

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 20, 2006, 19:14:26 ; Search time 134 Seconds

(without alignments)  
2167.383 Million cell updates/sec

Title: US-09-463-733-1

Sequence: 1 MDENAIKAIPIQWYRRHQ.....VHHDIDPTCESKVIDPKKS 661

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A\_Geneseq.21:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*  
9: geneseqp2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3315	96.3	637	4	ABB71489	Abb71489 Drosophila
2	1305	37.9	722	8	ADN22998	Adn22998 Bacterial
3	1271.5	36.9	653	4	ABA47250	Ab47250 Human Pp7
4	1271.5	36.9	653	8	ADT6673	Adt6673 Human pro
5	1164.5	33.8	613	7	ADJ95096	Adj95096 Novel NOV
6	1164.5	33.8	613	7	ADJ95098	Adj95098 Novel NOV
7	581	16.9	488	8	ADY05391	Ady05391 Plant ful
8	577	16.8	520	4	ABB63840	Abb63840 Drosophila
9	571	16.6	500	4	ABG09989	Abg09989 Novel hum
10	570	16.6	494	5	ABG70122	Abg70122 Human pre
11	570	16.6	499	4	AAE09722	Aae09722 Novel cel
12	570	16.6	499	7	ADB64103	Adb64103 Human Pro
13	570	16.6	499	8	ABM82127	Abm82127 Tumour-as
14	569	16.5	499	7	ADB64101	Adb64101 Rat Prote
15	562.5	16.3	482	5	AAE16244	Aae16244 Tomato in
16	562.5	16.3	485	5	AAE16236	Aae16236 Tomato in
17	562.5	16.3	499	4	AAE09733	Aae09733 Protein p
18	560	16.3	499	4	AAE09732	Aae09732 Protein p
19	556	15.1	499	4	ADN21438	Adn21438 Bacterial
20	519	14.6	598	8	ABG09988	Abg09988 Novel hum
21	502.5	14.4	513	6	ABR53882	AbR53882 Protein s
22	496.5	13.9	513	7	ADK64862	Adk64862 Disease t
23	479.5	13.9	513	7	ADK64862	Adk64862 Disease t
24	479.5	13.9	513	7	ADK64862	Adk64862 Disease t

25	479.5	13.9	513	8	ADS43550	Ad43550 Bacterial
26	460.5	13.4	346	4	ABR59732	Abbr59732 Drosophila
27	451	13.1	244	4	ADS23974	Ad23974 Bacterial
28	435.5	12.7	323	6	ABP98033	Abp98033 Protein p
29	434.5	12.6	319	8	ADK87721	Adk87721 Plant ful
30	434	12.6	318	3	AAE30062	Aae30062 Arabidops
31	434	12.6	318	8	ADN72465	Adn72465 Thale cre
32	431.5	12.5	380	8	ADY09284	Ady09284 Plant ful
33	430.5	12.5	325	8	ADY59840	Ady59840 Plant ful
34	430.5	12.5	334	8	ADY08660	Ady08660 Plant ful
35	430.5	12.5	334	8	ADY22333	Ady22333 Plant ful
36	430	12.5	304	3	AAE30063	Aae30063 Arabidops
37	429.5	12.5	325	8	ADY57338	Ady57338 Plant pol
38	429.5	12.5	326	8	ADY59253	Ady59253 Plant pol
39	429	12.5	338	8	ADY10803	Ady10803 Plant ful
40	429	12.5	338	8	ADY06935	Ady06935 Plant ful
41	427.5	12.4	316	8	ADN19921	Adn19921 Bacterial
42	425	12.3	401	8	ADN23944	Adn23944 Bacterial
43	424.5	12.3	312	3	AAE39549	Aae39549 Arabidops
44	423	12.3	290	8	ADK67181	Adk67181 Plant ful
45	423	12.3	313	8	ADK79431	Adk79431 Plant ful

## ALIGNMENTS

RESULT 1  
ID ABB71489 standard; protein; 637 AA.  
XX ABB71489;  
AC ABB71489;  
DT 26-MAR-2002 (first entry)  
DE Drosophila melanogaster polypeptide SEQ ID NO 41259.  
XX Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
OS Drosophila melanogaster.  
XX MO200171042-A2.  
XX 27-SEP-2001.  
XX 23-MAR-2001; 2001WO-US009231.  
XX 23-MAR-2000; 2000US-0191637P.  
XX 11-UTL-2000; 2000US-00614150.  
XX (PEKE ) PE CORP NY.  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX WPI, 2001-656860/75.  
XX N-PSDB; ABL15392.  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
XX genes from Drosophila and for elucidating cell signaling and cell-cell  
XX interactions.  
XX Disclosure, SEQ ID NO 41259; 21pp + Sequence Listing; English.  
XX The invention relates to an isolated nucleic acid detection reagent  
XX capable of detecting 1000 or more genes from Drosophila. The invention is  
XX useful in developmental biology and in elucidating cell signalling and  
XX cell-cell interactions in higher eukaryotes for the development of  
XX insecticides, therapeutics and pharmaceutical drugs. The invention  
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
XX sequences (ABL101840-ABL16175) and the encoded proteins (ABR57737-  
XX ABB72072). The sequence data for this patent did not form part of the  
XX printed specification, but was obtained in electronic format directly  
XX from WIPO at ftp.wipo.int/pub/published\_pct\_sequences



Oy	275	SPSTSLDLILKSIDRGVYVILAPPLTDSERPLDK-----TEQQQLPDLIMSPDPAKTMGC	327
Db	388	SDQTEVSKLDKIPRRHPOSILRPPVKKGMSESEKNSAVNVDEKQMLDLMSPDKONKGC	447
Oy	328	VPNTLRGAGVWGPDVTDFLORHRLSYVIRSHCEKCPNGHEFMDHKITITFSASNYVAI	387
Db	448	WPNVFRGGSGSYFGADITLAFLEKHGRLLVVRSHCEKFGEGYFSPHNNTCLTVPSASYYEF	507
Oy	388	GSNKAAYIR-LNNQLMPHFVQYITSAASQTRLS-FKQMGIVESSALKEIAYMRDHRDE	445
Db	508	GSNRGAAYVFIKSKQPHFVQYTM--ASKTRKSTLIERLGVVESSAVKEIKELSFPHND	565
Oy	446	LEDERKTKDPKQSGYISTSHMCKMENVNFKLGPYMLLDKLAAP-CTDSQKVVNNTTDL	504
Db	566	LOKEBIMDIKSGKLPILKMSDCERRITGLMLPWALAPKATLSBDEKLYVVKEDRRI	625
Oy	505	LDTVTLLEAADMGSVMALYANKASLVAFNIIADNDSGEITLDEFPETALIDLVAHMPG	564
Db	626	AQVGSTHAQEKD---IVESLYNHHKSTLETLPFRMDNDNGQVSMKEFIDACEVL-----G	677
Oy	565	AYSKAEM-----LEKCRMDLNGDGKVDLINEFLAEFRLSD	599
Db	678	KYTKRPLQTDYISQIAESIDFNMDDGFIDNLELAEFRLVD	717
RESULT 3			
ID	AA847250		
AC	AA847250	standard; protein; 653 AA.	
XX	AA847250;		
DT	18-JUL-2001	(first entry)	
XX	Human PP7.		
DE			
KW	Mitochondrial calcium-binding protein; p95.6/YN52; calcium;		
KV	drug screening; sequence extraction; PP7; calcium binding motif.		
OS	Homo sapiens.		
XX			
PN	WO200130830-A2.		
PD	03-MAY-2001.		
XX			
PF	26-OCT-2000; 2000MO-US029787.		
XX			
PR	26-OCT-1999; 99US-00427867.		
PR	26-OCT-1999; 99US-00427895.		
PR	07-SEP-2000; 2000US-00657253.		
XX			
PA	(MITO-) MITOKOR.		
XX			
PI	Murphy AN, Fahy ED, Wiley SE, Clevenger W;		
XX			
DR	WPI: 2001-308616/32.		
DR	N-PSDB; AAC85804.		
XX			
PT	New Caenorhabditis elegans mitochondrial calcium-binding protein,		
PT	p95.6/YN52, identified by protein motif database searching, useful for		
XX	regulating intracellular calcium ions and in drug screening assays.		
PS	Example 2; Fig 13; 176pp; English.		
XX			
CC	This sequence shows a protein containing calmodulin/calcium binding		
CC	motifs which was identified using the method of the invention. The method		
CC	comprises a computer system for extracting information from biological		
CC	sequence databases, which receives a motif that defines a desired		
CC	polypeptide sequence. The computer searches the database based on a query		
CC	formulated from the motif to obtain a set of polypeptide sequences which		
CC	are then filtered and a report of the filtered set of polypeptide		
CC	sequence information is then displayed. The computer system is useful for		
CC	identifying recombinant C.elegans p95.6/YN52 and other mitochondrial		

protein involved in binding, transport or other regulation of intracellular calcium which provides targets for screening assays that identify agents capable of altering a mitochondrial activity

Sequence 653 AA:

Query Match	36.9%	Score 1271.5	DB 4	Length 653
Best Local Similarity	40.3%	Pred. No. 1.2e-110		
Matches 258	Conservative 129	Mismatches 204	Indels 49	Gaps 11
QY	3	ENAIRAIFIOKWRBRHQAAREMORRCNMOIFQNLVYASRQDAQLEKFPNDLIKHHMQA	62	
DB	14	DTSLAAALIIINMWRGKRLAKARQYALITFOSIEYADEQOQQLTFSSFMLENTHTI	73	
QY	63	AGRNKOYOGSAHVVLDDKD--DLVE--BFQDIYNA-KIELPIRKNIIDLLIDVFRKRG	117	
DB	74	HKEBELRLNQSLBSEQDMRDMVDVDSIDVPDSYNGSRRLQEPFLCTDIDLLBFRKQO-	132	
QY	118	NRLHPKTYALLREAAKSLKQLEPNISPVSTAVSQOQTVCCGDHGKLDLVLVHKGLPS	177	
DB	133	-ILHAHYLVLEVPETKVKLKQMPNPTIIGTSPSKERTVTCGLHKKLDLFLIPKNGLPS	191	
QY	178	SSNYPVNGDVPDVGKGLVTLTLLSLYLAPAAVPLANGNHDSTMAARXGIRBEVS	237	
DB	192	ERNPVPNGDVPDVGKNSIBILMLCVFLVYPNDLHANGNHDSTMAARXGIRBEVS	251	
QY	238	KYPRNHRKILAFIDEVVRWMLPLGSLVLSRVLYVIGGFSGSTSLDILKSIDRGKVSILRP	297	
DB	252	KYKLGKRIQLIIEFPFAMLPITGIVNDELIVHGGISFTDMLLKRVERANKKSLIIP	311	
QY	298	PL-----TDGEP---LDKTEWQIIPDIMSDPQATYGCV	328	
DB	312	PTEINRBDHTDSKKNKGVTPNAHGRILKTNGSPTEHLTBHMEWQIIDILMSDPKNGQCF	371	
QY	329	PNTLRGAGWGFQPVDTNFLQRRRLSVYISHSHECKPGHEPMDNKIITTFPSASNTYALG	388	
DB	372	PNTRGGGCYGPGPVTSKILNKYQKMLIRSHSECKPGGYELCHDGKVVITFSAANYLEG	431	
QY	389	SNKGAVYIRLNQMLPHFVQYISASQTRKLSF---KQRMGVSSALKEILAVRRDRDE	445	
DB	432	SNKGAVYIRLCSGTTTFRFQY---QVTKATCFQRLRQVDTMESAIIKILBERTSRKSD	487	
QY	446	LEDEFRRKYDPKDSGYISISHMKVEMENVTKLGLPWRLLRDLKAPGDSOKVNVNRTIDL	505	
DB	488	LTRAFQGLDHRKSKGLSVSQWAFCEBNILGLNTPMRSLSNLVNIIDQNGVEYSSPQNI	547	
QY	506	DTD-VILAEADGMSVMDALYANKASLVALFNITDADNGSEITLDEBFTALIDLVAHMPG	564	
DB	548	RIEKPVOEASH---TLVETTLRYRSDEIIIFNADIDTHSGILSVBFRAMWKLFFSSHYNV	604	
QY	565	AYSKAKMLEKCRMDNDGDKVDLNEPTELEAFRLSDLRKE	604	
DB	605	HIDDSQVKNKLINIMDLNKGSDIDNEPLKAPYV--VARYE	642	
RESULT 4				
ADT66673				
ID	ADT66673	standard; protein; 653 AA.		
XX	ADT66673;			
AC	16-DEC-2004	(first entry)		
DT				
XX				
DE				
XX				
XX				
KW		Human protein phosphatase, EF hands-1 protein.		
KW		pain regulation; screening; Lu2P; PBP-19; phosphatidylinositol synthase;		
KW		valosin; interleukin 6 receptor subunit beta; aspartate aminotransferase;		
KW		neuronal immediate early gene; heat shock protein 27; HSC70; calmodulin;		
KW		synactin binding protein 1; epliceosoma protein SAP 155; neurodop 1;		
KW		bamacan; leukotriene A4 hydroxylase; chondromodulin 1;		
KW		Mse4/phosphatidylinositol-4-phosphate 5 kinase; 26S proteasome UB p112;		
KW		UB Z; UB p44.5; ingensin; 3-hydroxy-3-methylglutaryl-Co A synthase;		
KW		phosphoglycerate kinase; RNA polymerase II TF SIII p18 UB;		

KM protein phosphatase EF hands-1; RAB21;  
 KM TNF receptor-associated factor 6-binding protein; spaghetti protein;  
 KM HSKM-B; calnexin-C; nexin 3; glia-derived nexin precursor;  
 KM microtubulin-associated protein 4; CDC10;  
 KM transcription co-activator CRSP150; JERKY; SOUL protein;  
 KM rRNA intron-encoded homing endonuclease;  
 KM epithelial cell growth inhibitor; glutathione transferase A4; CGI-69;  
 KM differential display; chronic pain; gene therapy.  
 XX  
 OS Homo sapiens.  
 PN MO2004070383-A2.  
 XX  
 PD 19-AUG-2004.  
 XX  
 PF 02-DEC-2003; 2003WO-BP013597.  
 XX  
 PR 05-FEB-2003; 2003DB-01004680.  
 XX  
 PA (CHER ) GRUNENTHAL GMBH.  
 PI Weihe E, Bieller A, Schaefer MK;  
 XX MPI, 2004-604492/58.  
 DR N-PSDB; ADT66672.  
 XX  
 PT New polynucleotides and polypeptides involved in pain regulation, useful  
 XX in screening for pain-regulating compounds for therapy and diagnosis.  
 PS Claim 1; Fig 42b; 241pp; German.  
 XX  
 CC This invention describes novel polynucleotides and polypeptides involved  
 CC in pain regulation which are useful in screening for pain-regulating  
 CC compounds for therapy and diagnosis. Identifying pain-regulating  
 CC substances comprises incubating a test compound with a cell, and/or  
 CC preparation of a cell, that synthesises at least one specific peptide or  
 CC protein, then measuring either binding of test compound to the protein or  
 CC a functional parameter that is altered by the binding. The protein is one  
 CC or more of any of luzp, PEP-19, phosphatidylinositol synthase, valosin-  
 CC containing protein, interleukin 6 receptor subunit beta, aspartate  
 CC aminotransferase, neuronal immediate early gene, heat shock protein 27,  
 CC HSC70, calmodulin, syntaxin binding protein 1, spliceosomal protein SAP  
 CC 155, neuropilin 1, bamacan, leukotriene A4 hydrolase, chondromodulin 1  
 CC Mas4/phosphatidylinositol-4-phosphate 5 kinase, 26S proteasome UB p112 or  
 CC UE 2 or UE p44.5, ingensin, 3-hydroxy-3-methylglutaryl-Co A synthase,  
 CC phosphoglycerate kinase, RNA polymerase II TF SIII p18 UB, protein  
 CC phosphatase EF hands-1, RAB21, TNF receptor-associated factor 6-binding  
 CC protein, ortholog of the Drosophila gene 'spaghetti', protein with 4-  
 CC transmembrane domains member 3, vacuolar protein sorting protein, HSKM-B,  
 CC calnexin -C, sorting nexin 3, glia-derived nexin precursor, microtubuli-  
 CC associated protein 4, CDC 10, transcription co-activator CRSP150, JERKY,  
 CC XM-209578, TF SOX10, SOUL protein, rRNA intron-encoded homing  
 CC endonuclease, epithelial cell growth inhibitor, glutathione transferase  
 CC A4 and/or CGI-69. The specified nucleic acids and proteins have been  
 CC identified using the restriction fragment differential display PCR method  
 CC (Biochem. Biophys. Res. Commun.: 234 (1997) 16), applied to rats in which  
 CC arthritis has been induced by injection of Complete Freund's adjuvant.  
 CC Compounds identified by the new method, also specified polynucleotides,  
 CC their antisense sequences, encoded proteins and peptides, vectors,  
 CC antibodies and cells, are useful as therapeutic, and diagnostic agents,  
 CC particularly for treatment of chronic pain, including by in vivo or in  
 CC vitro gene therapy. This sequence represents the human protein  
 XX  
 XX  
 SO Sequence 653 AA;  
 Query Match 36.9%; Score 1271.5; DB 8; Length 653;  
 Best Local Similarity 40.3%; Pred. No. 1.2e-110;  
 Matches 258; Conservative 129; Mismatches 204; Indels 49; Gaps 11;  
 QY 3 ENAIRAALPFOKYRRHQARRRQRCNNQIFONTLYASQDQAEIYKFFNDLIKMPQA 62  
 DB 14 DTSIRALALIONMYRGYKALKRKAROHVALTIPQISIEVADEQGMQISTFSPLENTYTHI 73

QY 63 AGRKNQYOGSAHVSYLDDKD--DLVE--BFGDIYVA-KIEPIPRKNDILIDVPRKKRG 117  
 DB 74 HKEBELKNSLESQDRDRNDYDSIDVPSYNGPKLQFPLCTDIDLILBAKEQG- 132  
 QY 118 NKLHPRKYALLIRBAKSLIKQLPNISPVSTAVSQOVTYCGDLHGKLDLLVYLKNGLPS 177  
 DB 133 -ILHAHYVLEVLFEKTKVLKQMPNFTHTIQTSPSKSVTICGDHLGKLDLFLIFYKNGLPS 191  
 QY 178 SGNPVYVGGDFVDRKRGLEVTLLLSLYLAFPMNAVFLNRGNHDSVNAARGFIREVDS 237  
 DB 192 KRNPFVFGDFVDRKNGSLIIMILCVSFLVYPNDLILNRGNHDFPMNLRYGFTRKILH 251  
 QY 238 KYPRNHKQILAFIDVYRWLPFGSVLANSRVLTVHGAFSPDSTLDLKSIDRGKYYSILRP 297  
 DB 252 KYKLHGKRIQLQLEBFYAMLPVIGTIVDEILVYHGISBTIDNLILHVRKNGMSVLIP 311  
 QY 298 PL-----TDGERP--LDKTEWQIIFDIWMSDPOATMGCV 328  
 DB 312 PTEINRDHDTDSKNNKVGVTFNAHGRIKTNGSPTEHLTSEHWEQIIDLMSDPGRKNGCF 371  
 QY 329 PNTLAGAGVWRGPDYTDNPLQHRSLSVYRSHECKPNHGFHGDKILITIFASNYAAG 388  
 DB 372 PNTCRGGGCFEPDVTSKILNKYQLKMLIRSHCKPBGVEICHDKKVVITFASNYBERG 431  
 QY 389 SNKGAVIRLNQMLMPFVOYISAASQTKRLSF---KQKGIYESSALKELAVRMDHRDE 445  
 DB 432 SNRGAVITLCGSGTTRFPQY---QVTKATCQPLRQVNDTMENAIKILRRERVSRSKD 487  
 QY 446 LEDFRKYPDKDSGYISISHWCKMENVTGLPWRLLRDKLAPQTDQKQVNYNRTDLL 505  
 DB 488 LTRAFQLDHRKSGKLSVQMAFCMENITGLNLPKRSLSNINVDQNGNVEYMSFQNI 547  
 QY 506 DTD-VILREADGMSVMDALYNKASLYAIFNIDADNGSETLDEFEPAIDLLVAHMG 564  
 DB 548 RIERSKQVGRHS--TLVEVLVYRSDLEIIFPAIDTDSGLSVBERFRAMTLFESHYNV 604  
 QY 565 AVSKALEKCKMDINGKQVDLNFLEAFPLSDLRKE 604  
 DB 605 HLDSDQVKLANIMDLNKGDSIDPNEFLKAFV--VHRYE 642  
 RESULT 5  
 ADJ95096  
 ID ADJ95096 standard; protein; 613 AA.  
 AC ADJ95096;  
 DT 06-MAY-2004 (first entry)  
 XX  
 DB Novel NOVX protein sequence #162.  
 XX  
 XX antidiabetic; anorectic; cardiant; hypotensive; antiarteriosclerotic;  
 XX anorectic; virucide; antibacterial; fungicide; protozoacide; nootropic;  
 XX neuroprotective; antiparkinsonian; anticonvulsant; osteopathic;  
 XX antiarthritic; antiinflammatory; dermatological; antiaslthmatica;  
 XX antilipemic; gene therapy; metabolic disorder; diabetes; obesity;  
 XX infectious disease; anorexia; cancer; cardiovascular disease;  
 XX hypertension; atherosclerosis; neurodegenerative disorder;  
 XX Alzheimer's disease; Parkinson's disease; epilepsy; immune disorder;  
 XX osteoarthritis; hematopoietic disorder; inflammatory skin disorder;  
 XX asthma; dyslipidemia; neurogenesis; cell differentiation;  
 XX cell proliferation; hematopoiesis; wound healing; angiogenesis;  
 XX chromosome mapping; tissue typing; pharmacogenomic.  
 OS Homo sapiens.  
 PN MO2003040325-A2.  
 XX  
 PD 15-MAY-2003.  
 XX  
 PF 05-NOV-2002; 2002WO-US035464.  
 XX



PR	05-NOV-2001,	2001US-0338656P.
PR	06-NOV-2001,	2001US-03330722P.
PR	09-NOV-2001,	2001US-0348283P.
PR	15-NOV-2001,	2001US-0335610P.
PR	16-NOV-2001,	2001US-0338543P.
PR	20-NOV-2001,	2001US-0331630P.
PR	20-NOV-2001,	2001US-0331641P.
PR	21-NOV-2001,	2001US-0332152P.
PR	27-NOV-2001,	2001US-0333461P.
PR	28-NOV-2001,	2001US-0333912P.
PR	28-NOV-2001,	2001US-0334027P.
PR	29-NOV-2001,	2001US-0334300P.
PR	30-NOV-2001,	2001US-0334421P.
PR	30-NOV-2001,	2001US-0334526P.
PR	04-DEC-2001,	2001US-0336576P.
PR	04-DEC-2001,	2001US-0336664P.
PR	07-DEC-2001,	2001US-0338314P.
PR	07-DEC-2001,	2001US-0338390P.
PR	10-DEC-2001,	2001US-0339006P.
PR	10-DEC-2001,	2001US-0339008P.
PR	11-DEC-2001,	2001US-0339286P.
PR	01-FEB-2002,	2002US-033380P.
PR	01-FEB-2002,	2002US-0353288P.
PR	04-FEB-2002,	2002US-0354392P.
PR	04-FEB-2002,	2002US-0354393P.
PR	04-FEB-2002,	2002US-0354409P.
PR	27-FEB-2002,	2002US-0359944P.
PR	27-FEB-2002,	2002US-0360148P.
PR	05-MAR-2002,	2002US-0361790P.
PR	05-MAR-2002,	2002US-0361833P.
PR	05-MAR-2002,	2002US-0361925P.
PR	05-MAR-2002,	2002US-0362230P.
PR	05-MAR-2002,	2002US-0362655P.
PR	13-MAR-2002,	2002US-0364000P.
PR	13-MAR-2002,	2002US-0364181P.
PR	13-MAR-2002,	2002US-0364182P.
PR	13-MAR-2002,	2002US-0364197P.
PR	13-MAR-2002,	2002US-0364227P.
PR	17-MAY-2002,	2002US-0381621P.
PR	28-MAY-2002,	2002US-0383675P.
PR	17-JUL-2002,	2002US-0396703P.
PR	06-AUG-2002,	2002US-0401552P.
PR	07-AUG-2002,	2002US-0401594P.
PR	07-AUG-2002,	2002US-0401787P.
PR	15-AUG-2002,	2002US-0403619P.
PR	20-AUG-2002,	2002US-0404821P.
PR	23-AUG-2002,	2002US-0405688P.
PR	23-AUG-2002,	2002US-0405402P.
PR	23-AUG-2002,	2002US-0405496P.
PR	23-AUG-2002,	2002US-0405631P.
PR	26-AUG-2002,	2002US-0406125P.
PR	04-NOV-2002,	2002US-00287226.

CC The invention relates to novel isolated polypeptides, mature forms of  
CC these, or a sequence that is at least 95 % identical to, or having one or  
CC more conservative amino acid substitutions in the polypeptides. The  
CC polypeptides, nucleic acid molecules and antibodies are useful in the  
CC manufacture of a medicament for treating a syndrome associated with a  
CC human disease, preferably a NOVX-associated disorder. The nucleic acid  
CC molecules, polypeptides and antibodies are useful for treating,  
CC preventing or diagnosing diseases such metabolic disorder, diabetes,  
CC obesity, infectious diseases (viral, bacterial, fungal, helminthic, and  
CC protozoal), anorexia, cancer, cardiovascular diseases (hypertension,  
CC atherosclerosis), neurodegenerative disorders, Alzheimer's disease,  
CC Parkinson's disease, epilepsy, immune disorders (osteoarthritis),  
CC hematopoietic disorders, inflammatory skin disorders, asthma, and various  
CC dyslipidemias. The nucleic acids and polypeptides may also be used as  
CC targets for the identification of small molecules that modulate or  
CC inhibit e.g. neurogenesis, cell differentiation, cell proliferation,  
CC hematopoiesis, wound healing and angiogenesis, in gene therapy, in  
CC generation of antibodies that bind immunospecifically to NOVX substances  
CC for use in therapeutic or diagnostic methods. The nucleic acids are  
CC further used as hybridization probes, in chromosome mapping, tissue  
CC typing, preventive medicine, and pharmacogenomics. This sequence  
CC corresponds to one of the NOVX polypeptides of the invention.  
SQ Sequence 613 AA;

Query Match	33.8%	Score 1164.5	DB 7	Length 613
Best Local Similarity	39.3%	Pred. No. 1.5e-100		
Matches	247	Conservative 115	Mismatches 200	Indels 67
			Gaps 13	
QY	3	ENALDAIFIOKVRHROARREMOECWMOIFONLEIYASQDAIBLYKEFNDLIKMPQA	62	
DB	14	DTSLFAALLIIONWYGYARLAKAROHVALLTIFOSIEIYADECOQOLSTFFSEFLBRYTHI	73	
QY	63	AGRKQYOGSAHVSLDDKD--DYVE--BPGIVNA-KIELEIRKXHDILLDFVFKKGG	117	
DB	74	HKEBELRNQSIIESQDMRDWDYDSIDVPDSYNGPRQIFPLTCTDIDILLBAFSGQ-	132	
QY	118	NRLHEKYVALILREAAKSLKOLPNISPVSTAVSQOVTYCGDLHGKLDLLVYLNKGLPS	177	
DB	133	-ILNHHVLYLVLFETKVKYLKQMPNFTHITQTSPSKEVTICGDLHGKLDLFLIFYKGLPS	191	
QY	178	GSNPFYVNGDPVDRKRGLEVLILLLSLYLAPNAVFLNRGNHDSVMNARVGFIREVS	237	
DB	192	ERNPFVFGDPDRDRKNSIEILMLCVSFLVYPNDLHNRGNHDEPMNLARGFTTBILH	251	
QY	238	KYPRHKKIILAFIDEVYRMLP-----LGSVLNRYLIVHGSPSDSTSLD	282	
DB	252	KYKLGKRIILQILBEFYAMLPEYTRKDGTSKHKXGVTFNA-----HGRK-----	299	
QY	283	IKSIDRGKYVILREPLTDGAB--LDKTEWQOIFDINMSDPOATWGCVPNTLRGAVWF	339	
DB	300	-----TNGSPTEHLTSHBMEQIIDLMSDPRGNKGCPPNTRCGGGCYF	342	
QY	340	GPDYTDNLEQHRHLSYVLSRSHCKNGCHEFMDNKLITTFSSANYYALGSNNGAYIRLNN	399	
DB	343	GPDVYSKILNKYQMLQMLIRSHCKEKGEGYICHDGKVITTFSSANYYIESBSNNGAYIKLOS	402	
QY	400	QLMPHFVYIISAQOTKRLSF---KQRMGIVESSALKEILAVMRDHRDLBEPFRKYDEK	456	
DB	403	GTTPRFPQY----QVTKATCFQPLRQVDTMENSAMIKILREKVVISKKSPLTFAFQODHR	458	
QY	457	DSGYISISHWKWENVTKGLPFRLLDKLAPGDSQKVNTRYRLTDLDTD-VIIIAEA	515	
DB	459	KSGKLSVQWAFPCMENIIGLNLPMRSLTSSNLVNIQNGNVEVWSSQNRIRKPYGEAHS	518	
QY	516	DGMSVMDLLYANKASIVAILFNIITDADNGSEITLDEFEETAILDLVHAMPGAYKAELEKQ	575	
DB	519	---TLVETLYRYRSLDEIIFNAIDVDHSGLISVEEFRAMMKLPSHHYNNHIDDSQYNKLA	575	
QY	576	RMMDLNGDKVDLNEFLAFLSLDKRKE	604	
DB	576	NIMDLNRGSDIDFNEFLKAFVY--VHRYE	602	







Qy	121	HPKVAIIILAEAAKSLKOLPNISEVSTAASQOVTWCGDHLGDLDDLVLVHLKNGLPSSSN	180
Db	209	HRKCAVYLIVQVKEVLSKSLSTVETTLKSTETKTLVCGDTHQGFYDLMLIFELNGLPSESTN	266
Qy	181	PVYVNGDFVYDGRKRGLEVLNLLSLVLAEPNAVFLIRKGNHDSVNNAAVYGFIREVSEKYP	240
Db	269	PYLINQGFVYDGRSFSEVILLTLRGFLLLRPDHNLIRGNHRLDNNQIYGFIRGVRKAKYT	322
Qy	241	RNHKRIIAFLDEVYRWLPGLSVINSRVLLVHGQF--SDSTSLDLIKSIDRGKYVSIILRP	296
Db	329	---AQWYELSFSEVPEMLPLAQCIINGKVLIMHGGLFSEDDVTLDDIRKIERNR-----QRP	380
Qy	299	LTOSEERLDKTEWMOQIIDIWMSDQALTMCGCVPTMLGAGVWREPRDVTDFLQRHLSYVIR	350
Db	381	--DSGF-----CDLWMSDQOPONG--RSISKGGVTCQFPRDVTKAFIERNNMLDYIIR	422
Qy	359	SHECKRPFHEFMHDKNKLITTFPSASNYAIGSNKAGAYRL--NQOLMPEHFOY	408
Db	430	SHEVKAGEVYVAHGRCVTFSPANVCDOMKAKASTIHLQSSDLNRPQHOFR	480
RESULT 10			
ABG70122			
ID	ABG70122	standard; protein; 494 AA.	
AC			
XX	ABG70122;		
XX			
DT	21-OCT-2002	(first entry)	
XX			
DE		Human prey protein for Shigella ospC1 #40.	
XX			
KW	Prey protein; ospB; ospD1; ipaB; ipaC; ipaH; 8; ospG; ospC1; Shigella;		
KW	Shigellois; bacillary dysentery; antibacterial; yeast two-hybrid system.		
XX	protein-protein interaction; SID; selected interacting domain; human.		
XX			
OS	Homo sapiens.		
XX			
XX	Wo200257303-A2.		
XX			
PN	25-JUL-2002.		
PD			
XX			
PE	11-JAN-2002; 2002MO-EP000777.		
XX			
PR	12-JAN-2001; 2001US-0261130P.		
XX			
PA	(HYBR-) HYBRIGENICS.		
XX			
PI	Legrain P;		
XX			
XX	WPI; 2002-599706/64.		
DR	N-PSDB; ABS51515.		
XX			
XX			
PT	New complex of protein-protein interactions between a bait Shigella		
FT	flexneri polypeptide and a prey mammalian or human placenta polypeptide		
PT	for treating or preventing bacillary dysentery in a mammal or human.		
XX			
PS	Claim 7; Page 90-91; 162pp; English.		
XX			
CC	The invention relates to a complex of protein-protein interactions		
CC	between a Shigella flexneri polypeptide (e.g. ospB, ospD, ipaB, ipaC,		
CC	ipaH, 8, ospG and ospC1) and a mammalian polypeptide defined in the		
CC	specification. The complexes are formed using the yeast two-hybrid		
CC	system. Also included are (1) a recombinant host cell expressing the		
CC	interactions between the Shigella flexneri polypeptide and a mammalian		
CC	polypeptide defined in the specification; (2) selecting a modulating		
CC	compound that inhibits or activates the protein-protein interactions;		
CC	(3) a modulating compound obtained from the method of (2); (4) a SID		
CC	(selected interacting domain) polypeptide or its fragment or variant		
CC	comprising the human polypeptides appearing as ABG70042-ABG70242; (5) a		
CC	SID polynucleotide or its fragment or variant comprising encoding the		
CC	above polypeptides a vector comprising (5); (6) a recombinant host cell		
CC	containing the vector; and (10) a protein chip comprising Shigella		
CC	flexneri polypeptide and a mammalian polypeptide defined in the		

CC		specification. A pharmaceutical composition comprising the compound,
CC		polypeptide or polynucleotide is useful for treating or preventing
CC		shigellosis (bacillary dysentery) in a human or mammal. The present
CC		sequence represents a human prey protein isolated by the yeast two-hybrid
CC		assay, forming a complex of the invention with a shigella protein
XX		
SQ	Sequence 494 AA;	
Query Match	16.6%; Score 570; DB 5; Length 494;	
Best Local Similarity	33.7%; Pred. No. 2,3e-44;	
Matches 143; Conservative	71; Mismatches 142; Indels 68; Gaps 16	
QY	11 FICGKRRHQAAREMQRCKMQIQONLEAYEEDQDAELKYFPNDLIKMPQAAGRKNQY 70	
DB	90 YIKGYRRAAASNMALGK-----FR---AALRD-----YEIVKKYPHDKAKMKYQ 132	
OY	71 -----GSAAV-SVLDDKD-----DLVEBFG--DIVNAKIELPIRKNNIDL 107	
DB	133 ECKKIYKOKAPERRAIAGDEHKRSVVDSLIDISMTEDEDYSQPKLEDGKVTFISFKE---- 188	
OY	108 LIDVFKRKRGNRLHPKYVALILREAAKSILKQLPNISPSTAVSQCVTVCGDLAGKLDLL 167	
DB	189 LMQWYKQKG-KLHRKCAVOYLIVQKVELSLSTLVETTLTKETKITVCGDTHQGPYDLL 246	
OY	168 VILHKNGLPSSSNPFYENGDFVDNRGKLBYALLLSLYLAFRNAVFLNRGNHDSVMNA 227	
DB	247 NIFELNGLPSETNTPEYIIFMGDFVDRGSPFEVEYILLTLEGKLIYPPDHFIILRGNHETDNMQ 306	
OY	228 RYGRIIRVESKPYPRNHNKRIIAFDIEVYWNLPLGSTLNRYLIVHGCF--SDSTSILDKS 285	
DB	307 IYGEGBGVKAKYT--AQMYELFSBEVFEMWLPAQCINRNVLMHGRLPSBDGVTLDDLRK 363	
OY	286 IDRKSUYSLARPPLTDGEPRLDKTEWOQIFDIIMSDPQATMGCVPTLTLRGAGVWFGPDVT 345	
DB	364 IERNR-----QRP--DSGPM-----CDLMSDPQONG-RSISKRGVSQGFPGDYTK 407	
OY	346 NFLOPRHLSTYIRSHCEKCPNGHEFMENDKITTPISASYVAIGSKKAYIRL-INQALMFH 404	
DB	408 AFLKENNIIDYTIIRSHBVAAEGYEVAAHGRCVTVFESAPMYCDQMGNKASYIHILGSDLRFQ 467	
OY	405 FVOY 408	
DB	468 FHQP 471	
RESULT 11		
AAE09722		
ID	AAE09722 standard; protein; 499 AA.	
XX		
AC	AAE09722;	
XX		
DT	29-NOV-2001 (first entry)	
XX		
DE	Novel cell cycle protein, protein phosphatase type 5 (PP5).	
XX		
KW	Cell cycle protein; DNA repair; protein phosphatase type 5; PP5;	
KX	gene therapy; cancer; vaccine; cytostatic; vulnertary; antiinflammatory.	
OS	Unidentified.	
XX		
FH	Key Location/Qualifiers	
FT	Misc-difference 1..7	
FT	/note= "Encoded by CGAG; This translation exception	
FT	occurs while decoding with AAD16789"	
XX		
FN	MO200164913-A2.	
PD	07-SEP-2001.	
XX		
PF	02-MAR-2001; 2001MO-USO06849.	
XX		
PR	02-MAR-2000; 2000US-00517779.	
XX		



```

QY 71 -----GSAHV-SVLDKD-----DLVEBFG--DIVNAKTELPFRKHIDL 107
DB 138 ECKKIYKQAKAFERAIADDEKRSVVDLIESMTIEDYSGPKLEDKVTISFMKE---- 193
QY 108 LIDVFRKRGNRRLAPKYVALILREAKSLKQLEPNISPVSTAVSQOVTVCGLDGLKLDL 167
DB 194 LMQWYDQK--KLHRCAYQILVQVKEVLSKSTLVBTELLKEBKITVCGDTHGQPYDL 251
QY 168 VILHKNGLPSSSNPYVNGDFVDRGKGLSVLLLSLYAFPAVNAVFLNKGHEDSVMA 227
DB 252 NIFELNGLPSETMYPYIFNGDFVDRGSPSVSEVILTLFGRKLYDPHFLLRGNHETDMNQ 311
QY 228 RYGFIREVSKYPRNHRKILAFIDEYVRMLPLGSVLSRVLIVHGGF--SDSTSLDLIKS 285
DB 312 IYFEGEVKAKYT---AQWTELFSEVFEWLPLOCTINGKVLNMGGLPSEGVTLDIRK 368
QY 286 IDRGKYVSLRPPLTDEGPLDKTEWQOIFDIMSDPOATWGCVPNTLRGAGVFGPVD 345
DB 369 IERNR-----QPF--DSGPM-----CDLMSDPQPNG-RSISKRGVSCQPGPVT 412
QY 346 NPLQHRRLSVYIRSHCKPENGHEFMHDKITITFSASNTYAIGSNKGAAYIRL-NNOLMPH 404
DB 413 AFLERNNLDIYIIRSHVKAEGYEVANHGRCVTFSPAPNYCDQGNKASYIHLQSDLRPQ 472
QY 405 FVOY 408
DB 473 FHQF 476

RESULT 13
ID AEM82127 standard; protein; 499 AA.
XX AEM82127;
DT 18-NOV-2004 (first entry)
XX
DE Tumour-associated antigenic target (TAT) polypeptide PRO25881, SEQ:5498.
XX
KM Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KM colorectal cancer; cell proliferative disorder; breast cancer;
KM colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KM central nervous system cancer; bladder cancer; pancreatic cancer;
KM cervical cancer; melanoma; leukaemia; hybridisation probe;
KM chromosome identification; chromosome mapping; gene mapping;
KM gene therapy; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2004030615-A2.
XX
PD 15-APR-2004.
XX
PF 29-SEP-2003; 2003WO-US028547.
XX
PR 02-OCT-2002; 2002US-0414971P.
XX
PA (GETH ) GENENTECH INC.
XX
PI Wu TD, Zhang Z, Zhou Y;
XX
DR WPI, 2004-347921/32.
XX
DR N-PSDB; ACN40614.
XX
PT New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.
XX
PS Claim 12; SEQ ID NO 5498; 7273bp; English.
XX
CC The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are

```

```

CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT polypeptide of the invention
XX
SQ Sequence 499 AA;
Query Match 16.6%; Score 570; DB 8; Length 499;
Best Local Similarity 33.7%; Pred. No. 2.4e-44;
Matches 143; Conservative 71; Mismatches 142; Indels 68; Gaps 16;
QY 11 FIOKTRRHQAREMORRCMWOIFONLRYASEDDQALEYFNDLIKMPQAAGRKQYQ 70
DB 95 YKGYVYRRAASNNALCK-----FR-----AALRD-----YETVVKVXPHDXDAKKYQ 137
QY 71 -----GSAHV-SVLDKD-----DLVEBFG--DIVNAKTELPFRKHIDL 107
DB 138 ECKKIYKQAKAFERAIADDEKRSVVDLIESMTIEDYSGPKLEDKVTISFMKE---- 193
QY 108 LIDVFRKRGNRRLAPKYVALILREAKSLKQLEPNISPVSTAVSQOVTVCGLDGLKLDL 167
DB 194 LMQWYDQK--KLHRCAYQILVQVKEVLSKSTLVBTELLKEBKITVCGDTHGQPYDL 251
QY 168 VILHKNGLPSSSNPYVNGDFVDRGKGLSVLLLSLYAFPAVNAVFLNKGHEDSVMA 227
DB 252 NIFELNGLPSETMYPYIFNGDFVDRGSPSVSEVILTLFGRKLYDPHFLLRGNHETDMNQ 311
QY 228 RYGFIREVSKYPRNHRKILAFIDEYVRMLPLGSVLSRVLIVHGGF--SDSTSLDLIKS 285
DB 312 IYFEGEVKAKYT---AQWTELFSEVFEWLPLOCTINGKVLNMGGLPSEGVTLDIRK 368
QY 286 IDRGKYVSLRPPLTDEGPLDKTEWQOIFDIMSDPOATWGCVPNTLRGAGVFGPVD 345
DB 369 IERNR-----QPF--DSGPM-----CDLMSDPQPNG-RSISKRGVSCQPGPVT 412
QY 405 FVOY 408
DB 473 FHQF 476

RESULT 14
ID ADE64101 standard; protein; 499 AA.
XX ADE64101;
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein P53042, SEQ ID NO 10049.
XX
KM Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KM chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS Rattus norvegicus.

```







XX Claim 16; Page; 75pp; English.  
 PS The patent discloses cell cycle proteins and nucleic acids encoding them.  
 CC The cell cycle proteins are capable of promoting cell cycle progression  
 CC or cell cycle arrest. They are capable of promoting or inhibiting DNA  
 CC repair. The invention also provides methods for screening a bioactive  
 CC agent capable of modulating the cell cycle or interfering with the  
 CC binding of a cell cycle protein, protein phosphatase type 5 (PP5). The  
 CC method involves combining a cell cycle protein and a candidate bioactive  
 CC agent and determining the effect on the cell in the presence or absence  
 CC of the candidate agent. Cell cycle proteins bind to Rad9 proteins. They  
 CC are useful for generating an agonist and antagonist of the cell cycle  
 CC proteins, both of which are useful in gene therapy and in screen assays.  
 CC They are useful in the diagnosis and treatment of cancer. The proteins of  
 CC the invention are used as vaccines. They are useful for generating  
 CC polyclonal or monoclonal antibodies to the proteins. The present sequence  
 CC is N303A variant of a novel cell cycle protein, protein phosphatase type  
 CC 5 (PP5) of the invention. Note: The present sequence is not shown in the  
 CC specification, but is derived from the PP5 protein shown in Fig 2  
 CC (AA809722)  
 XX

SO Sequence 499 AA;

Query Match 16.3%; Score 562; DB 4; Length 499;

Best Local Similarity 33.5%; Pred. No. 1.4e-43;

Matches 142; Conservative 71; Mismatches 143; Indels 68; Gaps 16;

QY 11 FLOKMYRRHQAAREMQRRCNMOIFQNLRYASBODAEYKFFNDLIKMPQAAGRNQY 70  
 DB 95 YIKGYRRRAASNMALGK-----FR-----AALRD-----YETVVKVPHDKDAMKQY 137  
 QY 71 -----GSAHV-SVLDDKD-----DLYEERG--DIYAKIELPIRKNIID 107  
 DB 138 ECKNIVKOKAFERAIAGDEHKRSVVDLSISMTIDEYSGPRLBQKTTISPMKE----- 193  
 QY 108 LIDVFRKRGKGNLHPYVALILREBAKSLKQLPNISPVSTAVSQVTVCGDLHGKLDLL 167  
 DB 194 LMQWYKDKQ--KLHRCAYQILVQVKEVLSKSTLVETTLKTEKITVCGDTHGQFYDIL 251  
 QY 168 VLIHNGRLPSSSNPYFVNDPVDGRGRLGLVLLLSLTVLAPNNAVFLRGNHSDSVMA 227  
 DB 252 NIFELNGLPSEETNPYIFNDPVDGRGFSVEVILTLFGFLVLPDHFHLRGNHJETDNNMQ 311  
 QY 228 RYGFIREVESKYPKRNHKLIAFIDVYRWLPGLSVLNSRVLVHGGF--SDSTSLDLIS 285  
 DB 312 IYFGSGEVAKYKT--AQMTELFSEVFEWLPQAOCINGKVLIMHGGLFSEBDVTLDDIRK 368  
 QY 286 IDRGKYVSLRPPLTGDSEPLDKTEWQIFDIMMSDPOATWGCVPNTLRGAGWFGPDVTD 345  
 DB 369 IERNR-----QPF--DSGPM-----CDLMSDPOQNG--RSISKRGVSCQFGPDVTK 412  
 QY 346 NPLQGRRLSYVTRSHCKRNGHFPMDNKIITIFASNTYALGSKNGATYRL--NNQDMH 404  
 DB 413 AFLRENNLLYILIRSHKAVGKGVNAHGRCVTFSAPNYCDQMGNTASYIHLQGSDLRQ 472  
 QY 405 FVQY 408  
 DB 473 FHQF 476

RESULT 18

AA809731

ID AA809731 standard; protein; 499 AA.

XX AA809731;

XX 29-NOV-2001 (firec entry)

XX Protein phosphatase type 5 (PP5) variant, H244A.

DB Cell cycle protein; DNA repair; protein phosphatase type 5; PP5;

XX gene therapy; cancer; vaccine; cytostatic; vulnertary; antiinflammatory;

KW mutant; mutein; variant.

XX Unidentified.

OS Key Location/Qualifiers

FT Misc-difference 244 /note="wild-type His is substituted with Ala"

PN MO200164913-A2.

PD 07-SEP-2001.

XX 02-MAR-2001; 2001MO-US006849.

PF 02-MAR-2000; 2000US-00517779.

PR (RIGB-) RIGEL PHARM INC.

XX Xu X, Luo Y;

XX MPI; 2001-557775/62.

PT Cell cycle protein having phosphatase activity for generating agonist and

PT antagonist e.g. antibodies of the polypeptides, both of which are useful

PS in gene therapy and in screen assays.

XX Claim 16; Page; 75pp; English.

XX The patent discloses cell cycle proteins and nucleic acids encoding them.

CC The cell cycle proteins are capable of promoting cell cycle progression

CC or cell cycle arrest. They are capable of promoting or inhibiting DNA

CC repair. The invention also provides methods for screening a bioactive

CC agent capable of modulating the cell cycle or interfering with the

CC binding of a cell cycle protein, protein phosphatase type 5 (PP5). The

CC method involves combining a cell cycle protein and a candidate bioactive

CC agent and determining the effect on the cell in the presence or absence

CC of the candidate agent. Cell cycle proteins bind to Rad9 proteins. They

CC are useful for generating an agonist and antagonist of the cell cycle

CC proteins, both of which are useful in gene therapy and in screen assays.

CC They are useful in the diagnosis and treatment of cancer. The proteins of

CC the invention are used as vaccines. They are useful for generating

CC polyclonal or monoclonal antibodies to the proteins. The present sequence

CC is H244A variant of a novel cell cycle protein, protein phosphatase type

CC 5 (PP5) of the invention. Note: The present sequence is not shown in the

CC specification, but is derived from the PP5 protein shown in Fig 2

CC (AA809722)

XX

SO Sequence 499 AA;

Query Match 16.3%; Score 560; DB 4; Length 499;

Best Local Similarity 33.5%; Pred. No. 2.1e-43;

Matches 142; Conservative 71; Mismatches 143; Indels 68; Gaps 16;

QY 11 FLOKMYRRHQAAREMQRRCNMOIFQNLRYASBODAEYKFFNDLIKMPQAAGRNQY 70

DB 95 YIKGYRRRAASNMALGK-----FR-----AALRD-----YETVVKVPHDKDAMKQY 137

QY 71 -----GSAHV-SVLDDKD-----DLYEERG--DIYAKIELPIRKNIID 107

DB 138 ECKNIVKOKAFERAIAGDEHKRSVVDLSISMTIDEYSGPRLBQKTTISPMKE----- 193

QY 108 LIDVFRKRGKGNLHPYVALILREBAKSLKQLPNISPVSTAVSQVTVCGDLHGKLDLL 167

DB 194 LMQWYKDKQ--KLHRCAYQILVQVKEVLSKSTLVETTLKTEKITVCGDTHGQFYDIL 251

QY 168 VLIHNGRLPSSSNPYFVNDPVDGRGRLGLVLLLSLTVLAPNNAVFLRGNHSDSVMA 227

DB 252 NIFELNGLPSEETNPYIFNDPVDGRGFSVEVILTLFGFLVLPDHFHLRGNHJETDNNMQ 311

QY 228 RYGFIREVESKYPKRNHKLIAFIDVYRWLPGLSVLNSRVLVHGGF--SDSTSLDLIS 285

DB 312 IYFGSGEVAKYKT--AQMTELFSEVFEWLPQAOCINGKVLIMHGGLFSEBDVTLDDIRK 368





```

QY 72 SAHVSVLDDKCD---DLVE-----BFGIVNAKIELPFRKHNDILLID-----11.0
Db 190 -----DNKDAKTLKVECEKIVKQLAFPAALIEVGBELSAAGLDVESMAVDASYDGVRL 24.2
QY 111 -----VFRKKGNRLHPKRYVALILREBAKSLKOLPNISPVSTAIVCOQVTCG 15.7
Db 243 EGNEMTQEFIDDMIERFFKKGKLIHKKYVQIILAVANIVYNEPVTWVEVDIPEDVQILVCG 30.2
QY 158 DLHG-----KDDLVLVHLKNGLPSSSNPPYVFNEDPFDYGRKRGLE 19.7
Db 303 DTHGFMALAMPSPSXWLLTNCEXAQYFDMELERLNLGFPEDKRYTILFNEDFVDRGSWSTE 36.2
QY 198 VLLILLSTLYLAPPNNAVFLNRGNHSDVMNARVGFPIREVSKY-----23.9
Db 363 IALLLYAYKMLARPNGFFIRNGNHETDMMRVYGFEBEGCEKHNIVBRASRTLCHAYATSK1 42.2
QY 240 -----PRNHKRILAFIDVYRWMLPLGVSILNSRYLVHVGCF--SDSTSLDLIKS 28.5
Db 423 FEHTPADTSHLLPRITYK----LFSBSPSLPLATLIGKFKFLVHGGJLFSDDNVTLLDIRK 47.8
QY 286 IDRGKVSILRBPPLTDGEPDLDKTEWQOITPDIMASDPOATMGCPYNTLRGAGVFFGPDVTD 34.5
Db 479 LDRHQ-----KOPGOAG-----LMEMEMLTWDQOPFGGRPSK-RGVGMQFGPDVTK 52.4
QY 346 NLOHRRLSYVLRSHCEKNGHBFMEDNKILITIFSAASYAIAISNGAYVIRLNNOLMHPH 40.5
Db 525 RPCDNGLEAIRSHHVAMDGYBEHRDGKCIYFSA.PKCDMTEKNGAYINIGPDYKLKP 58.4
QY 406 VOY 408
Db 585 SQF 587

```

CC	The invention relates to isolated polynucleotide (I) and polypeptide (II).
XX	
PS	Claim 20; SEQ ID NO 40347; 103bp; English.
XX	
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
XX	biodiversity.
XX	
DR	N-PSDB; AAS74175.
XX	
PI	Drmancac RT, Liu C, Tang YT;
XX	
PR	31-MAR-2000; 2000US-00540217.
XX	
PA	(HYSR-) HYSEQ INC.
XX	
PF	30-MAR-2001; 2001WO-US006631.
XX	
PD	11-OCT-2001.
XX	
PN	WO200175067-A2.
XX	
OS	Homo sapiens.
XX	
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM	food supplement; medical imaging; diagnostic; genetic disorder.
XX	
DE	Novel human diagnostic protein #9979.
XX	
DT	13-FEB-2002 (first entry)
XX	
ID	ABG09988 standard; protein; 714 AA.
XX	
AC	ABG09988;
XX	
RESULT 22	
ABG09988	

CC reaction (PCR) primers, oligonucleotides, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
CC  
CC Sequence 714 AA;

Query Match	14.4%	Score 496.5	DB 4	Length 714
Best Local Similarity	29.2%	Pred. No. 3.9e-37		
Match 140	Conservative 62	Mismatches 147	Indels 131	Gaps 15
QY	SEDOAALYKFEFNNLIKMPQAARKQYQ-----	GSAY-SYLDSD	82	
DB	211 SEHEAATLYLL---VVKYKPHDKAKKKYDGCNKIVAKAPERALAGDEHGRSVVDSIDI	287		
QY	83 ---DLVEBFQ--DIVNAKIELPIRKNHIDLLIDVFRKKGRNLHPKYVALILREAKSLK	137		
DB	288 ESMATIEDBSGPKLEDGKVTISFMKE---LMQWYKQK--KLARKCAVQILVQKVEVLS	341		
QY	138 QLPNISPSTAVSQWYVCGEDLHKLDLLVVLKNGLPSSNPVY-----	183		
DB	342 KLSLTVETTLTKETKIVCGDTHQPYDLLNI FELINGLPSETNPVVSFAEPPSPHLHQ	401		
QY	184 -----FNGDPVDRGKRGLEVLNLLLSLYLAF	209		
DB	402 PQPQVCTWEHPEPLPKQVTLERRLAMLHGQGLFNGDPVDRGSPSVSEVILITLFGKILLY	461		
QY	210 PNAVFLNRGNHEDSVNARVGFIREVSKYPRNHKRIILAFIDEVYRMPLGSGVINSRYLI	269		
DB	462 PDHHLTLRGNHETDNMNIQVGFEGEVAKYTL---AQWELPSVFEWLPPLAQINGKYLII	518		
QY	270 VHGF--SDTSLLISIDRGKAVSLRPLLTGSEPLDTEKMOQIDPIMSDQ-----	322		
DB	519 MHGGLFEBDGTLLTDIKIRERNR-----QRP--DSALPRRMGFVPLNAGSGSPQFPHLR	571		
QY	323 ---ATMGCVPTLL-----	RGAGVWGPEDVTDNPLQ	349	
DB	572 LGPVHLHGAAMASTVVTWVRGSGTVAENTCPYLFNGRSISKGVSCQFSPDVTAKALE	631		
QY	350 RHRLSYVIRSHCEKPNHGHEFMDKRIITIFASNYVAIGSNKGAIVRL--NRQLMPHFVQY	408		
DB	632 ENMIDIIRSHSEVVAEGEVAHNGRCVTVFSAPRYCOMGNKASYIHLQSSDLRQPHQF	691		
RESULT 23				
ABRS53882	ABRS53882			
ID	ABRS53882 standard; protein; 513 AA.			
AC	ABRS53882;			
XX				
DT	20-JUN-2003 (first entry)			
XX				
DE	Protein sequence #SEQ ID 2629.			
XX				
XX	Multiprotein complex; eukaryote; drug target; diagnosis.			
OS	Saccharomycetes cerevisiae.			
XX				

PN EPI258494-A1.  
 XX  
 PD 20-NOV-2002.  
 XX  
 XX 20-DEC-2001; 2001BP-00130253.  
 PP  
 XX 15-MAY-2001; 2001BP-00111774.  
 PR  
 XX (CELL-) CELLZOME AG.  
 XX  
 PA Bauer A, Gavin A, Grandi P, Krause R, Kruse UD, Kuester BD,  
 PI Marzloch M, Schultz JD, Superti-Furga GD;  
 XX MPI; 2003-250078/25.  
 DR N-PSDB; ACC61924.  
 XX  
 XX New isolated protein complexes useful for diagnosing a disease or  
 PT disorder, or as a target for an active agent of a pharmaceutical,  
 PT preferably a drug target in the treatment or prevention of disease or  
 PT disorder.  
 XX  
 PS Disclosure; SEQ ID NO 2629; 17pp + Sequence Listing; English.  
 XX  
 XX The invention relates to multiprotein complexes from eukaryotes. Proteins  
 CC of the invention and DNA sequences encoding them are given in records  
 CC ABR52568-ABR53903 and ACC60610-ACC61944 respectively. The complexes are  
 CC obtainable by using a protein as a bait and isolating the set of proteins  
 CC which is attached thereto from cells. Such protein complexes may comprise  
 CC up to 30 distinct proteins. Protein complexes of the invention are useful  
 CC for diagnosing a disease or disorder, or as a target for an active agent  
 CC of a pharmaceutical, preferably a drug target in the treatment or  
 CC prevention of a disease or disorder. Note: The sequence data for this  
 CC patent is not represented in the printed specification, but is based on  
 CC sequence information supplied by the European Patent Office. The complete  
 CC document is available on CD-ROM  
 XX  
 XX Sequence 513 AA;  
 SQ  
 Query Match 13.9%; Score 479.5; DB 6; Length 513;  
 Best Local Similarity 32.3%; Pred. No. 9.3e-36;  
 Matches 130; Conservative 68; Mismatches 154; Indels 51; Gaps 13;  
 QY 2 DENAIAAIFIOKMYRRHQAARREM--ORRCMIOFONLEVASBODQAEUYKFNLDLHG 58  
 DB 113 DPAATKALITCRFIRBERFRKAIGAKENAKISLCQTNLSSFPDANADLANY----- 165  
 QY 59 MPOAAGRKOYOGSAHVSLDDKDLVEBFQDIVNAKIELPIRKNHIDLIDVFRKRGK 118  
 DB 166 ---BGPKEFE-----QLYDDKN-----AFKGAIKRMQGEFISKVNDLFLK--GK 207  
 QY 119 RLHPKYVALILREAAKSLKQLPNISPV--STAVSQQVTCGDLGKLDLVLVHLKNGLP 176  
 DB 208 YLPKTYVAALISHADTLFRQEPSEWELNNSTPDVKISVCGDTHQGFYDLVNLFRKFGKV 267  
 QY 177 SSNPFVFGDFVDRKRGLEVLILLISYLAFPNAPVFLNRGNHSDVNNARVGFIREVE 236  
 DB 268 GPKHTYLFNGDFYDRGSMSCVALLFYCKILHPNPFILNRGNHSDNNKIKYGFEDCK 327  
 QY 237 SKYPRNKKIILAFIDEVYKRLPLGASVLSKRVLIHGCF--SD--STSLDLKSIDRGKTVSI 294  
 DB 328 YKTL---SQRIFNNFASFSPLPLATLINDYLVMHGGLPSDPSATLSDPKNIDR-----F 379  
 QY 295 LRPPLTDEBPLDKTIEWQOIFDIMMSDPOATMGCVPTLIGAGVMPGPVDTNDFLQHRHS 354  
 DB 380 ADPP--RDG-----ATMELIMADPOEANGMGPSQ--RGLGAPFDPDITDRFLNNKLR 428  
 QY 355 YVIRSHCKPENGHEFMHDKNITITFSASNYTLAGSKGAYIRL 397  
 DB 429 KIRSHSLHNGKGVGFQKQKGLMTVFSAPIYCSQSGNLGVYIH 471  
 RESULT 24  
 ADK64862

ID ADK64862 strand; protein; 513 AA.  
 XX  
 AC ADK64862;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DB Disease treating protein complex-derived protein #1585.  
 XX  
 XX protein complex; drug target; diagnosis.  
 XX  
 OS Unidentified.  
 XX  
 PN EPI338608-A2.  
 XX  
 XX 27-AUG-2003.  
 PD  
 XX 20-DEC-2002; 2002BP-00102902.  
 XX  
 XX 20-DEC-2001; 2001BP-00130253.  
 PR  
 XX (CELL-) CELLZOME AG.  
 PA  
 PI Bauer A, Gavin A, Superti-Furga G, Kuester B, Schultz J,  
 PI Marzloch M, Grandi P, Krause R, Kruse U, Merlino A, Bauch A,  
 PI Michon A, Leutwein C, Rick J;  
 XX MPI; 2003-638460/61.  
 DR N-PSDB; ADK64863.  
 XX  
 XX New proteins and protein complexes from eukaryotes, useful as targets in  
 PT drug screening, or in diagnosing or screening for the presence of a  
 PT disease or disorder, or a predisposition for developing a disease or  
 PT disorder in a subject.  
 XX  
 PS Disclosure; SEQ ID NO 3169; 13pp; English.  
 XX  
 XX The invention relates to novel protein complexes comprising a first and a  
 CC second protein, or its derivative, fragment, homologue or variant. The  
 CC proteins are selected from given protein complexes, which are not defined  
 CC in the specification. The variants are encoded by nucleic acids that  
 CC hybridize to the nucleic acids encoding the proteins under low stringency  
 CC conditions. The protein complexes are useful as targets for an active  
 CC agent of a pharmaceutical. These protein complexes are particularly  
 CC useful as drug targets for the treatment or preventing of a disease or  
 CC disorder. The complexes and methods above are useful in diagnosing or  
 CC screening for the presence of a disease or disorder or a predisposition  
 CC for developing a disease or disorder in a subject. These are also useful  
 CC in screening for a drug for treatment or prevention of a disease or  
 CC disorder. The molecule that modulates the amount, activity or protein  
 CC components of the complex is useful for the manufacture of a medicament  
 CC for the treatment or prevention of a disease or disorder. This sequence  
 CC corresponds to a protein of the invention. (Note: the sequence data for  
 CC this patent did not form part of the printed specification but was  
 CC obtained from the EPO in electronic format).  
 XX  
 XX Sequence 513 AA;  
 SQ  
 Query Match 13.9%; Score 479.5; DB 7; Length 513;  
 Best Local Similarity 32.3%; Pred. No. 9.3e-36;  
 Matches 130; Conservative 68; Mismatches 154; Indels 51; Gaps 13;  
 QY 2 DENAIAAIFIOKMYRRHQAARREM--ORRCMIOFONLEVASBODQAEUYKFNLDLHG 58  
 DB 113 DPAATKALITCRFIRBERFRKAIGAKENAKISLCQTNLSSFPDANADLANY----- 165  
 QY 59 MPOAAGRKOYOGSAHVSLDDKDLVEBFQDIVNAKIELPIRKNHIDLIDVFRKRGK 118  
 DB 166 ---BGPKEFE-----QLYDDKN-----AFKGAIKRMQGEFISKVNDLFLK--GK 207  
 QY 119 RLHPKYVALILREAAKSLKQLPNISPV--STAVSQQVTCGDLGKLDLVLVHLKNGLP 176  
 DB 208 YLPKTYVAALISHADTLFRQEPSEWELNNSTPDVKISVCGDTHQGFYDLVNLFRKFGKV 267







Db 222 IGPELKLEYQVFAVRPH 238

RESULT 28  
ABP98033  
ID ABP98033 standard; protein; 323 AA.

AC	ABP98033;
XX	
DT	11-AUG-2003 (first entry)

DE Protein phosphatase stress-related polypeptide GmPP2A-4.

KM Protein phosphatase stress-related polypeptide; PPSR; PpPP2A-1; PpPP2A-2; BnpPP2A-1; BnpPP2A-2; BnpPP2A-3; GmPP2A-1; GmPP2A-2; GmPP2A-3; OsPP2A-1; OsPP2A-2; OsPP2A-3; plant tolerance; environmental stress; plant; drought; salinity; cold; enzyme.

OS Glycine max.

PN WO2003020914-A2.

PD 13-MAR-2003.

PF 05-SEP-2002; 2002WO-US028445.

PR 05-SEP-2001; 2001US-0317305P.

PA (BADI ) BASF PLANT SCI GMBH.

PI Bohnert HJ, Chen R, Ishitani M, Van Thielen N, Da Costa E SilvaO;

DR WPI; 2003-300886/29.

DR N-PSDB; ACC43348.

PT New protein phosphatase stress-related polypeptide coding nucleic acid,  
PT useful for modulating plant's tolerance to an environmental stress such as  
PT as drought, increased salinity and cold.

PS Claim 13; Page 81; 107pp; English.

The present sequence represents a protein phosphatase stress-related polypeptide (PPSRP). The specification describes PPSR polypeptides designated PpPP2A-1, PpP-1, BmPP2A-1, BmPP2A-2, BmPP2A-3, GmPP2A-1, GmPP2A-2, GmPP2A-3, OsPP2A-1, OsPP2A-2 or OsPP2A-3, derived from *Physcomitrella patens*, *Brassica napus*, *Glycine max* or *Oryza sativa*. Polypeptides and polynucleotides are useful for modulating plant tolerance to an environmental stress such as drought or increased salinity and cold. They are also useful in identification and localization of *Physcomitrella patens*, *Brassica napus*, *Glycine max* or *Oryza sativa* and related organisms, mapping of genomes of organisms related to the above species, in a evolutionary and polypeptide structural studies, in determination of PPSR regions required for function, modulation of PPSR activity, modulation of metabolism of one or more cell functions and transmembrane transport of one or more components

**SQ Sequence 323 AA;**

Query Match	12.7%	Score 435.5	DB 6	Length 323
Best Local Similarly	31.2%	Pred. No. 6.4e-32		
Matches 104; Conservative	66;	Mismatches 124;	Indels 39;	Gaps 9;

[illegible]

```
Db
110 YPENEFLINGNHECASINRIYGFVDECKRRF--NRLMKTFITDCNCLPVAAIVDEKITL 166
```

```

QY      269  IYHGFS-DSTLIDLKSIDRGKVTSLRPPYLTGEPDLKTEWQOIFDINKMSDQATMG 327
      :|||:  :||:::  |  :  :  :|||:  :
Db      167  CMHGGLSPDLNINLDQIRLQF-----PTVDPTGLTCLDLMSPSKDVQG 211

```

QY 328 VPNTLRGAGVFGPDPVTDNPLQRRRLSYVRSHECPNGHFMNDKIITTPSASNYAI 387

Db 212 WGNDRGVSYTFGADKVSQFLQKDDLDVCRAHQVVEDGEYFFPANNQLVTTIFSPAPNYCGE 2712

```

QY      388 GSNKGAVIRLNQMLPHFVQYIYSAASQTRLSF 4200
      |||::|||::|||
Db      272 PDNAGAMMSVDETLMCSF-QILKPADKAKALNF 3033

```

RESULT 29  
ADX87721  
ID ADX87721 standard; protein; 339 AA

AC ADX87721;

DT 21-APR-2005 (first entry)

DE Plant full length insert polypeptide seqid 50385.

KM plant protectant; plant growth regulant; gene therapy; plant;  
KM recombinant DNA construct; physical array; plant breeding marker;  
KM cold tolerance; heat tolerance; drought tolerance; herbicide tolerance

growth rate; cell cycle pathway; disease resistance;

KW yield; plant growth; plant development; seed oil; protein yield;

[illegible]

05 Unidentified

PN US2004034888-A1

PD 19-FEB-2004.

PF 28-APR-2003

PR 06-MAY-1999; 99US-00304517
































PA (ZHOU/) ZHOU Y.

PA (SCRE/) SCREEN S E.  
PA (TABA/) TABASKA J E.  
PA (CAOY/) CAO Y.

PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JB, Cao Y, XX  
DR WPI; 2004-180133/17.

PT New recombinant DNA construct, useful for improving plant tolerance to PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or PT pests, for conferring increased resistance to plant disease, or for PT improving yield.

PS Claim 1; SEQ ID NO 50385; 15pp; English

CC The invention describes a recombinant DNA construct comprising a  
CC polynucleotide consisting of a sequence encoding an amino acid sequence  
CC available in electronic form from the US patent office at  
CC <http://seqdata.uspo.gov/sequence.html?docid:120043034888>. The polynucleotide  
CC of the invention are also useful in physical arrays of molecules and as  
CC plant breeding markers. The recombinant DNA construct is useful for  
CC improving plant tolerance to cold, heat, drought, herbicides, extreme  
CC osmotic conditions, pathogens or pests, for manipulating growth rate in  
CC plant cells by modification of the cell cycle pathway, for conferring



PR 19-JUL-1999; 99US-0144335P.  
 PR 20-JUL-1999; 99US-0144352P.  
 PR 20-JUL-1999; 99US-0144632P.  
 PR 20-JUL-1999; 99US-0144884P.  
 PR 21-JUL-1999; 99US-0144814P.  
 PR 21-JUL-1999; 99US-0145086P.  
 PR 22-JUL-1999; 99US-0145088P.  
 PR 22-JUL-1999; 99US-0145085P.  
 PR 22-JUL-1999; 99US-015087P.  
 PR 22-JUL-1999; 99US-015089P.  
 PR 22-JUL-1999; 99US-0145192P.  
 PR 23-JUL-1999; 99US-0145145P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 26-JUL-1999; 99US-0145227P.  
 PR 27-JUL-1999; 99US-0145913P.  
 PR 27-JUL-1999; 99US-0145918P.  
 PR 27-JUL-1999; 99US-0145919P.  
 PR 28-JUL-1999; 99US-0145951P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 03-AUG-1999; 99US-0146389P.  
 PR 04-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 05-AUG-1999; 99US-0147302P.  
 PR 05-AUG-1999; 99US-0147192P.  
 PR 06-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 09-AUG-1999; 99US-0147493P.  
 PR 09-AUG-1999; 99US-0147935P.  
 PR 10-AUG-1999; 99US-0148317P.  
 PR 11-AUG-1999; 99US-0148319P.  
 PR 12-AUG-1999; 99US-0148341P.  
 PR 13-AUG-1999; 99US-0148565P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149929P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 25-AUG-1999; 99US-0150566P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.

PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161922P.  
 PR 28-OCT-1999; 99US-0161933P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 12.6%; Score 434; DB 3; Length 318;  
 Best Local Similarity 29.6%; Pred. No. 8,7e-32;  
 Matches 106; Conservative 60; Mismatches 120; Indels 64; Gaps 11;

QY 59 MPQARKNQYQSAHVSVDDKDDVBERGDIIVNKKIEPIKKNHIDLDIVPRKR-- 116  
 1 MAEKPAOEQKQKAMEPAVID-----DIIRLRV-----FRNTRPG 36  
 DB 117 -GNRLHPKYVALIIRBAKSLKQLPNIS-----PVSTAVSOQVTCGDHAKLDILV 168  
 37 SGKQVH-----LSGG--EIRQLCAVSKRIFLQGNLBLEAPIKICGDIHQYSDLR 87  
 QY 169 VLRKNGLPSSSNPYVNGDFVDRGKRGLEYVLLLSLYLAFPAVPLNRGNHDSVMNAR 228  
 88 LPEYGGPPPEAN-YLFGIDVDGRKQSLFTICLLAVKIKYPNPFILRGNHESASINRI 146  
 QY 229 YGPIRVESEKYPNNHRIILAFIDEVYRWLPLGSVANSRLIVGGRS-DSTSLDKSID 287  
 147 YGFIYDECKRRF--NRLMKIFITDCNCLPVALIDRLICMGHGISPELKSIDQIRNIA 203  
 DB 288 RGYKVSILRPPLTDGSEPLDKTEWQIFDIMWSDPQATMGCPPTLRGAGVPGPDVTDNF 347  
 204 R-----PMDIPESGLVCDLIMSDPSGDVGKAND-RGVSTTFADKVAER 247  
 QY 348 LQRRRLSYVIRSHCEKNGHEFMHDKIITIFGASNYTALGSKGAYIRLNQJLMPHF 405  
 DB 248 LKRGHMDLICRAHQVVEDGYEFAEROLVTVFASPNYCGEFDNAGAMMSIDSLMCSF 305

RESULT 31  
 ADN72465  
 ID ADN72465 standard; protein; 318 AA.  
 XX  
 AC ADN72465;  
 XX  
 DT 15-JUL-2004 (first entry)  
 XX  
 XX Thale cress protein upregulated in E2Fa/Dpa expressing plants Segid 360.  
 XX  
 XX plant; transgenic; E2Fa/Dpa transcription factor; growth regulator;  
 KW animal feed product; thale cress; cell wall biosynthesis;  
 KW nitrogen metabolism; carbon metabolism.  
 OS Arabidopsis thaliana.  
 OS  
 PN WO2004035798-A2.  
 XX  
 PD 29-APR-2004.

XX 20-OCT-2003; 2003WO-BF011658.  
 XX PF  
 XX 18-OCT-2002; 2002EP-00079408.  
 XX PR  
 XX (CROP-) CROPDSSIGN NV.  
 XX PA  
 XX Inze D, De Veylder L, Vlieghe K;  
 XX PI  
 XX MPI; 2004-348466/32.  
 XX DR  
 XX N-PSDB; ADM72464.  
 XX DR  
 XX Altering plant characteristics, useful for producing plants for enzyme or  
 PT pharmaceutical production comprises modifying in a plant, expression of  
 PT one or more nucleic acids and/or modifying level or activity of one or  
 PT more proteins.  
 PS Claim 1; SEQ ID NO 360; 134pp; English.  
 XX  
 XX This invention relates to a novel method for altering one or more plant  
 CC characteristics. Specifically, it refers to identifying genes that are up-  
 CC - or down-regulated in transgenic plants overexpressing the heterodimeric  
 CC E2F4/Dpa transcription factor of Arabidopsis and using these sequences to  
 CC alter plant characteristics accordingly. The present invention describes  
 CC generating transgenic plants for the production of growth regulators,  
 CC enzymes, therapeutics, pharmaceuticals and animal feed products, where  
 CC the altered plant characteristics are selected from increased yield or  
 CC biomass, enhanced survival capacity, stress tolerance, plant architecture  
 CC or physiology, altered endoreduplication, biochemistry, signal  
 CC transduction, storage lipid mobilization and/or altered photosynthesis,  
 CC each relative to the corresponding wild type plants. Accordingly, these  
 CC sequences can also be useful as positive or negative selectable markers  
 CC during transformation of cells or tissues. The identified genes play a  
 CC role in a variety of biological processes such as DNA replication, cell  
 CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as  
 CC transcription factors. This polypeptide sequence is a cholesterin protein  
 CC expressed by a gene upregulated 1.3 fold or more in plants overexpressing  
 CC the E2F4/Dpa transcription factor, given in an exemplification of the  
 CC invention.  
 XX  
 XX Sequence 318 AA;  
 SQ  
 Query Match 12.6%; Score 434; DB 8; Length 318;  
 Best Local Similarity 29.6%; Pred. No. 8.7e-32;  
 Matches 106; Conservative 68; Mismatches 120; Indels 64; Gaps 11;  
 QY 59 MPQAGRKQYQGSAAHVSILDDKDLVEFGDVIWAKIELPIRKHIDLDIVFRKR-- 116  
 DB 1 MARKPAQBOBOKRAMBPAYLD-----DIIRRLV-----FRNTRRG 36  
 QY 117 -GRLRPKVALLLRBAASLKKLPNIS-----PVSTAVSQQVTVCCDHLKDDLLV 168  
 DB 37 SGQVH-----LSRG--EIRQLCAVSKEIFLQQLLELPKIKICGDIHQGYDLR 87  
 QY 169 VILKQNLPSSSNPFVNGDFVDRGKGLSVLLLSLYLPAVAFNKGHNSDVMNAR 228  
 DB 88 LFFYGFPPPAAN-YLPLGDIYVDKQSLKLTICLLAKIKYFNFLKGNHSAINNI 146  
 QY 229 YGFIRESKYPPNHRRIIAFDIVYRMPLGVSLSRVLIWGPS-DTSLDLIKSID 287  
 DB 147 YGFYDCKRFF--NVRLMKIFDPCFNCPLVALIDRIKCMGGISPLKSDQIRNTA 203  
 QY 288 RGRYVSLIRPPLDLDBLDTKTEWQIFDIWMSDPOATMGCVPTTLGAGVFPDYTDNF 347  
 DB 204 R-----PMDIPESGLVCDLMSDPSGVGWMND--GVSSTYFADVYASF 247  
 QY 348 LGRRLSYVIRSHCKNKGHPMNDKIITIFASNYVAIGSKGAYIRLNLQMLMHP 405  
 DB 248 LGRHMDLIRAGQVYVDGYEFPAERKQVTVFSAPIYCGFPAAGAMMSIDSLKCSF 305

RESULT 32  
 ADY09284

ID ADY09284 standard; protein; 380 AA.  
 XX  
 XX ADY09284;  
 AC  
 XX 21-APR-2005 (first entry)  
 XX DT  
 XX Plant full length insert polypeptide seqid 65099.  
 XX DE  
 XX plant protectant; plant growth regulator; gene therapy; plant;  
 KW recombinant DNA construct; physical array; plant breeding marker;  
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;  
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;  
 KW growth rate; cell cycle pathway; disease resistance;  
 KW galactomanan production; lignin production; plant growth regulator;  
 KW yield; plant growth; plant development; seed oil; protein yield;  
 KW protein content.  
 XX  
 XX Unidentified.  
 XX OS  
 XX US2004034888-A1.  
 XX PN  
 XX 19-FEB-2004.  
 XX PD  
 XX 28-APR-2003; 2003US-00425114.  
 XX PF  
 XX 05-MAY-1999; 99US-00304517.  
 XX PR  
 XX 05-NOV-2001; 2001US-00985678.  
 XX  
 XX (LIU/) LIU J.  
 XX PA (ZHOU/) ZHOU Y.  
 XX PA (KOVA/) KOVALIC D K.  
 XX PA (SCRE/) SCREEN S B.  
 XX PA (TABAR/) TABASKA J B.  
 XX PA (CAO/) CAO Y.  
 XX  
 XX Liu J, Zhou Y, Kovalic DK, Screen SB, Tabaska JB, Cao Y;  
 PI  
 DR MPI; 2004-180133/17.  
 XX  
 PT New recombinant DNA construct, useful for improving plant tolerance to  
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or  
 PT pests, for conferring increased resistance to plant disease, or for  
 PT improving yield.  
 XX  
 PS Claim 1; SEQ ID NO 65099; 15pp; English.  
 XX  
 XX The invention describes a recombinant DNA construct comprising a  
 CC polynucleotide consisting of a sequence encoding an amino acid sequence  
 CC available in electronic form from the US patent office at  
 CC ftp.segdata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide  
 CC of the invention are also useful in physical arrays of molecules and as  
 CC plant breeding markers. The recombinant DNA construct is useful for  
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme  
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in  
 CC plant cells by modification of the cell cycle pathway, for conferring  
 CC increased resistance to plant disease, for producing galactomanan,  
 CC lignin or plant growth regulators, for increasing the rate of homologous  
 CC recombination in plants, for improving yield by modification of  
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake  
 CC or by providing improved plant growth and development under at least one  
 CC stress condition or for modifying seed oil or protein yield and/or  
 CC content. This is the amino acid sequence of a plant full length insert  
 CC polypeptide that can be used in the recombinant DNA construct of the  
 CC invention.  
 XX  
 XX Sequence 380 AA;  
 SQ  
 Query Match 12.5%; Score 431.5; DB 8; Length 380;  
 Best Local Similarity 30.5%; Pred. No. 2e-31;  
 Matches 102; Conservative 66; Mismatches 125; Indels 41; Gaps 8;  
 QY 90 DIVNAKIELPIRKHIDLDIVFRKKRG--RLHPRYVALILREAAKSLKQPNISPVST 147

Db 62 DIINRLLEV-----RSRPGKQVQLSESRIRHLCAASRRIPLQGPVLELA 107

Qy 148 AVSQQVTCGDLHGKLDLVLVLRKNGLPSSSNPYVNGDFVDRGKGLVLLLSLYL 207

Db 108 PIR-----ICGVHGHQVSDLLRLPEYGGI.PPEAN-YLFGDYVDGKOSLFTICLLAVKI 162

Qy 208 AFPAVAVFLNRGNHDSVMNARVGFIRVESKYPNNHRILAFIDEVYRMPLGSLVNSRV 267

Db 163 KYPNFFPLRGNHHCASINRIYGYDECKRPF--NRLMKTPTECFNCPLVALIDKI 219

Qy 268 LIVHGFSS-DSTSLDLIKSIDRGKYVSLRPPLTDGSEPLDXTWQQIFDIMWSPDQATMG 326

Db 220 LCHHGGLSPDILNDQIRNLQR-----PTDVPDVGGLCDLILMSPSPSEVQ 264

Qy 327 CVPTLRGAGVWFGPDVTDNLFQHRRLSYVIRSHCKPENGHEFMNDKIIITIFASANYTA 386

Db 265 GWMGNDRGVSTYFGADKVSFELQKHDDLLICRAHQVVEDGVEFPANRQLVITFSAFNYCG 324

Qy 387 GSNKGAIVIRLNQMLMHPFVQYISAASQTKLSF 420

Db 325 EFDNAGMMWSVDETLMCSF-QILKPADKKAALNF 357

RESULT 33

ADTS9840

ID ADTS9840 standard; protein; 325 AA.

XX ADTS9840;

AC

DT 13-JAN-2005 (first entry)

XX

DE Plant polypeptide, SEQ ID 9917.

XX

KM Plant; transgenic; cold tolerance; growth rate; drought tolerance;

KM disease resistance; galactomannan production; plant growth regulator;

KM heat tolerance; herbicide tolerance; lignin production;

KM extreme osmotic condition tolerance; pathogens resistance;

KM pest resistance; yield improvement; seed oil yield; seed protein yield.

XX

OS Viaridiplantae.

PN US2004216190-A1.

XX

PD 28-OCT-2004.

XX

PF 18-DEC-2003; 2003US-00739930.

XX

PR 28-APR-2003; 2003US-00424599.

XX

PR 28-APR-2003; 2003US-00425115.

XX

PA (KOVA/) KOVALIC D K.

XX

PI Kovalic DK;

XX

PI WPI; 2004-757369/74.

DR

XX

PT New recombinant DNA constructs useful in the field of biochemistry and

PT genetics, and in particular for producing transgenic plants with improved

PT biological characteristics.

XX

PS Claim 2; SEQ ID NO 9917; 14pp; English.

XX

XX The invention relates a recombinant DNA construct comprising a

CC polynucleotide having any of 5544 nucleotide sequences (CDNAS SEQ ID NO:

CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences

CC (SEQ ID NO: 5545-11088). The CDNAS and proteins are from corn, soybean,

CC Arabidopsis, wheat and rape but the specification does not indicate which

CC sequences is derived from which organism. Also included is a method of

CC producing a plant having an improved property, comprising transforming a

CC plant with a recombinant DNA construct comprising a promoter region

CC functional in a plant cell operably joined to a polynucleotide encoding a

CC polypeptide associated with the property, and growing the transformed

CC plant. The property is selected from improving plant cold tolerance, for

CC manipulating growth rate in plant cells by modification of the cell cycle

CC pathway, for improving plant drought tolerance, for providing increased

CC resistance to plant disease, for galactomannan production, for production

CC of plant growth regulators, for improving plant heat tolerance, for

CC improving plant tolerance to herbicides, for increasing the rate of

CC homologous recombination in plants, for lignin production, for improving

CC plant tolerance to extreme osmotic conditions, for improving plant

CC tolerance to pathogens or pests, for yield improvement by modification of

CC photosynthesis, for modifying seed oil yield and/or content, for

CC modifying seed protein yield and/or content, for yield improvement by

CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake

CC and for yield improvement by providing improved plant growth and

CC development under at least one stress condition. The polynucleotide may

CC also encode a plant transcription factor. The methods and compositions of

CC the present invention are useful in the field of biochemistry and

CC genetics, in particular for producing transgenic plants with improved

CC biological characteristics such as increased yield, improved nitrogen

CC flow, increasing plant tolerance to cold or heat, improving plant

CC tolerance to extreme osmotic and drought conditions, and improving plant

CC tolerance to plant pests or pathogens. They can also be used in physical

CC arrays of molecules, plant breeding markers, computer-based storage and

CC analysis systems. The present sequence is one of the 5544 plant protein

CC sequences of the invention. Note: The sequence data for this patent did

CC not form part of the printed specification, but was obtained in

CC electronic format directly from USPTO at

CC seqdata.uspto.gov/sequence.html?docid=2040216190.

XX

SQ Sequence 325 AA;

Query Match 12.5%; Score 430.5; DB 8; Length 325;

Best Local Similarity 31.1%; Pred. No. 1.9e-31;

Matches 99; Conservative 69; Mismatches 122; Indels 29; Gaps 6;

Qy 89 GDIWNAKIEPLIRKNIHDLIDVFRKKGRNRLHPKVVALIRBAASLKLQPLNISPVSTA 148

Db 9 GAMEGALNDVYVR-----LVBSGRGRQVQLSEASIRQLCDVAKRVFLSQPML-----LR 59

Qy 149 VSOQVTCGDLHGKLDLVLVLRKNGLPSSSNPYVNGDFVDRGKGLVLLLSLYLA 208

Db 60 IQAPVVICGDIHQGFVLDLRLPLDGGVPPST-YVFLGDYVDGKOSLFTICLLAVKIR 118

Qy 209 FPAVAVFLNRGNHDSVMNARVGFIRVESKYPNNHRILAFIDEVYRMPLGSLVNSRV 268

Db 119 YPDKVPFLRGNHHCASINRIYGYDECKRPF--NRLMKTFSDCFNCPLFIALIDKIL 175

Qy 269 IVHGFSS-DSTSLDLIKSIDRGKYVSLRPPLTDGSEPLDXTWQQIFDIMWSPDQATMG 327

Db 176 CHHGGLSPDILNDQIRNLQR-----PAPFDPYGLCDLILMSPSPSEVQ 260

Qy 328 VPPTLRGAGVWFGPDVTDNLFQHRRLSYVIRSHCKPENGHEFMNDKIIITIFASANYTA 387

Db 221 WGSDDRGVSTYFGADKVSFELQKHDDLLICRAHQVVEDGVEFPANRQLVITFSAFNYCG 280

Qy 388 GSNKGAIVIRLNQMLMHPF 405

Db 281 FDNVGAALSIDENIMCSF 298

RESULT 34

ADY08660

ID ADY08660 standard; protein; 334 AA.

XX

AC ADY08660;

XX

DT 21-APR-2005 (first entry)

XX

DE Plant full length insert polypeptide seqid 64475.

XX

XX plant protectant; plant growth regulant; gene therapy; plant;

KM recombinant DNA construct; physical array; plant breeding marker;

KM cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;

KM extreme osmotic condition; pathogen tolerance; pest tolerance;

KM growth rate; cell cycle pathway; disease resistance;

KM galactomannan production; lignin production; plant growth regulator;  
 KM yield; plant growth; plant development; seed oil; protein yield;  
 KM protein content.  
 OS Unidentified.  
 XX  
 PN US2004034888-A1.  
 XX  
 PD 19-FEB-2004.  
 XX  
 PF 28-APR-2003; 2003US-00425114.  
 XX  
 PR 06-MAY-1999; 99US-00304517.  
 PR 05-NOV-2001; 2001US-00985678.  
 XX  
 PA (LIU/) LIU J.  
 PA (ZHOU/) ZHOU Y.  
 PA (KOVA/) KOVALIC D K.  
 PA (SCRE/) SCREEN S E.  
 PA (TABA/) TABASKA J E.  
 PA (CAOV/) CAO Y.  
 XX  
 PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y,  
 DR MPI, 2004-180133/17.  
 XX  
 PT New recombinant DNA construct, useful for improving plant tolerance to  
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or  
 PT pests, for conferring increased resistance to plant disease, or for  
 PT improving yield.  
 XX  
 PS Claim 1; SEQ ID NO 64475; 15pp; English.  
 XX  
 CC The invention describes a recombinant DNA construct comprising a  
 CC polynucleotide consisting of a sequence encoding an amino acid sequence  
 CC available in electronic form from the US patent office at  
 CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide  
 CC of the invention are also useful in physical arrays of molecules and as  
 CC plant breeding markers. The recombinant DNA construct is useful for  
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme  
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in  
 CC plant cells by modification of the cell cycle pathway, for conferring  
 CC increased resistance to plant disease, for producing galactomannan,  
 CC lignin or plant growth regulators, for increasing the rate of homologous  
 CC recombination in plants, for improving yield by modification of  
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake  
 CC or by providing improved plant growth and development under at least one  
 CC stress condition or for modifying seed oil or protein yield and/or  
 CC content. This is the amino acid sequence of a plant full length insert  
 CC polypeptide that can be used in the recombinant DNA construct of the  
 CC invention.  
 CC  
 XX  
 SO Sequence 334 AA:  
 Query Match 12.5%; Score 430.5; DB 8; Length 334;  
 Best Local Similarity 30.6%; Pred. No. 2e-31;  
 Matches 113; Conservative 67; Mismatches 134; Indels 55; Gaps 13;  
 QY 60 PQAAGKRYQSGAHYSVLDKDDLYBREFDIYNAK--IELPRKHIDLLIDYPRKKG 117  
 DB 16 PAAGG---QGGGGIDALL--DDIIRLLREVTARPGKQVQSSBIRQLCTVSR---- 65  
 QY 118 NRLHPKVALLIREAKSLKQLPNISPVSTAVSQVTVCGDLHAGKIDLLLVVLRKGLGS 177  
 DB 66 -----ALFLSQ-----PMLLELPKIR-----ICDHHQYSDLRLEPYGFP 105  
 QY 178 SSNPVYFNGDFVDRGKRGLEVLILLILSLYLAFFNAVFLNRKMGHEDSVNNAKYGFIREVS 237  
 DB 106 EAM-YIFLGDYVDRGKSGLETTICLLAYKIKYFENFLRKGNECASIRIYGFYBCKR 164  
 QY 238 KYRNRKRIALFIDERYRWLPPLGSLVNSRVLVHGGRS-DSTSLDLIKSIDRKYVSIIR 296  
 DB 165 RF---NVRIMKVFTEGFTLIPVALIDDKILCMHGGLSPDLAHLDBIRKNLQR----- 213

QY 297 PFLVDGEPLDKTEWQAIIPDIWMSDP-QATMGCVPTLIGAGVWFGPDYTDNFIQRRLSY 355  
 DB 214 -----PFDVDPQGLCCPLMSDPKQVQGGWMD-KVSYTFADKXSEFLQRHDL 265  
 QY 356 VIRSHCKPKNGHEFMDNKIITIFGASNYAIGSNKGYIRLNOLMPHFVQYISAASQT 415  
 DB 266 ICRAGQVVDGYEFPAFDRQLVTITFAPNVCGRFBDNAGMMSVDETLMSGF-QILMPAEK 324  
 QY 416 KRLSFRQRM 424  
 DB 325 QIYAKQNV 333  
 RESULT 35  
 ADY22333  
 ID ADY22333 standard; protein, 334 AA.  
 XX  
 AC ADY22333;  
 XX  
 DT 21-APR-2005 (first entry)  
 XX  
 DE Plant full length insert polypeptide seqid 70117.  
 XX  
 KM plant protectant; plant growth regulant; gene therapy; plant;  
 KM recombinant DNA construct; physical array; plant breeding marker;  
 KM cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;  
 KM extreme osmotic condition; pathogen tolerance; pest tolerance;  
 KM growth rate; cell cycle pathway; disease resistance;  
 KM galactomannan production; lignin production; plant growth regulator;  
 KM yield; plant growth; plant development; seed oil; protein yield;  
 KM protein content.  
 XX  
 OS Unidentified.  
 XX  
 PN US2004034888-A1.  
 XX  
 PD 19-FEB-2004.  
 XX  
 PF 28-APR-2003; 2003US-00425114.  
 XX  
 PR 06-MAY-1999; 99US-00304517.  
 PR 05-NOV-2001; 2001US-00985678.  
 XX  
 PA (LIU/) LIU J.  
 PA (ZHOU/) ZHOU Y.  
 PA (KOVA/) KOVALIC D K.  
 PA (SCRE/) SCREEN S E.  
 PA (TABA/) TABASKA J E.  
 PA (CAOV/) CAO Y.  
 XX  
 PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y,  
 DR MPI, 2004-180133/17.  
 XX  
 PT New recombinant DNA construct, useful for improving plant tolerance to  
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or  
 PT pests, for conferring increased resistance to plant disease, or for  
 PT improving yield.  
 XX  
 PS Claim 1; SEQ ID NO 70117; 15pp; English.  
 XX  
 CC The invention describes a recombinant DNA construct comprising a  
 CC polynucleotide consisting of a sequence encoding an amino acid sequence  
 CC available in electronic form from the US patent office at  
 CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide  
 CC of the invention are also useful in physical arrays of molecules and as  
 CC plant breeding markers. The recombinant DNA construct is useful for  
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme  
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in  
 CC plant cells by modification of the cell cycle pathway, for conferring  
 CC increased resistance to plant disease, for producing galactomannan,  
 CC lignin or plant growth regulators, for increasing the rate of homologous

PR	16-APR-1999,	99US-0128714F
PR	16-APR-1999,	99US-0129845P
PR	16-APR-1999,	99US-0130077P
PR	21-APR-1999,	99US-0130544P
PR	23-APR-1999,	99US-0130541P
PR	23-APR-1999,	99US-0130891P
PR	28-APR-1999,	99US-0131444P
PR	30-APR-1999,	99US-0132048P
PR	30-APR-1999,	99US-0132447P
PR	04-MAY-1999,	99US-0132448P
PR	05-MAY-1999,	99US-0132465P
PR	06-MAY-1999,	99US-0132467P
PR	07-MAY-1999,	99US-0132863P
PR	11-MAY-1999,	99US-0134218P
PR	14-MAY-1999,	99US-0134219P
PR	14-MAY-1999,	99US-0134221P
PR	14-MAY-1999,	99US-0134347P
PR	18-MAY-1999,	99US-0134768P
PR	19-MAY-1999,	99US-0134941P
PR	20-MAY-1999,	99US-0135124P
PR	21-MAY-1999,	99US-0135222P
PR	24-MAY-1999,	99US-0135353P
PR	25-MAY-1999,	99US-0136021P
PR	27-MAY-1999,	99US-0136392P
PR	28-MAY-1999,	99US-0136782P
PR	10-JUN-1999,	99US-0138540P
PR	10-JUN-1999,	99US-0138647P
PR	14-JUN-1999,	99US-0139131P
PR	16-JUN-1999,	99US-0139412P
PR	18-JUN-1999,	99US-0139453P
PR	17-JUN-1999,	99US-0139492P
PR	18-JUN-1999,	99US-0139454P
PR	18-JUN-1999,	99US-0139455P
PR	18-JUN-1999,	99US-0139461P
PR	18-JUN-1999,	99US-0139462P
PR	18-JUN-1999,	99US-0139463P
PR	18-JUN-1999,	99US-0139750P
PR	18-JUN-1999,	99US-0139763P
PR	21-JUN-1999,	99US-0139873P
PR	22-JUN-1999,	99US-0139889P
PR	23-JUN-1999,	99US-0140353P
PR	24-JUN-1999,	99US-0140354P
PR	24-JUN-1999,	99US-0140659P
PR	28-JUN-1999,	99US-0140823P
PR	29-JUN-1999,	99US-0140991P
PR	30-JUN-1999,	99US-0141287P
PR	01-JUL-1999,	99US-0141842P
PR	01-JUL-1999,	99US-0142154P
PR	02-JUL-1999,	99US-0142055P
PR	06-JUL-1999,	99US-0142330P
PR	08-JUL-1999,	99US-0142803P
PR	09-JUL-1999,	99US-0142937P
PR	12-JUL-1999,	99US-0142977P
PR	13-JUL-1999,	99US-0143542P
PR	14-JUL-1999,	99US-0143624P
PR	15-JUL-1999,	99US-0144005P
PR	16-JUL-1999,	99US-0144085P
PR	16-JUL-1999,	99US-0144086P
PR	19-JUL-1999,	99US-0144335P
PR	19-JUL-1999,	99US-0144331P
PR	19-JUL-1999,	99US-0144332P

PR 19-JUL-1999; 99US-014433P.  
 PR 19-JUL-1999; 99US-0144334P.  
 PR 19-JUL-1999; 99US-0144335P.  
 PR 20-JUL-1999; 99US-0144352P.  
 PR 20-JUL-1999; 99US-0144352P.  
 PR 20-JUL-1999; 99US-0144352P.  
 PR 21-JUL-1999; 99US-0144814P.  
 PR 21-JUL-1999; 99US-0145086P.  
 PR 21-JUL-1999; 99US-0145086P.  
 PR 22-JUL-1999; 99US-0145085P.  
 PR 22-JUL-1999; 99US-0145087P.  
 PR 22-JUL-1999; 99US-0145089P.  
 PR 22-JUL-1999; 99US-0145192P.  
 PR 23-JUL-1999; 99US-0145145P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 26-JUL-1999; 99US-0145276P.  
 PR 27-JUL-1999; 99US-0145313P.  
 PR 27-JUL-1999; 99US-0145318P.  
 PR 27-JUL-1999; 99US-0145319P.  
 PR 28-JUL-1999; 99US-0145351P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 02-AUG-1999; 99US-0146389P.  
 PR 03-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 04-AUG-1999; 99US-0147302P.  
 PR 05-AUG-1999; 99US-0147192P.  
 PR 05-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 09-AUG-1999; 99US-0147493P.  
 PR 09-AUG-1999; 99US-0147935P.  
 PR 10-AUG-1999; 99US-0148171P.  
 PR 11-AUG-1999; 99US-0148319P.  
 PR 12-AUG-1999; 99US-0148319P.  
 PR 13-AUG-1999; 99US-0148355P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 25-AUG-1999; 99US-0150566P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151388P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157533P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158332P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.

PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159337P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160982P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161928P.  
 PR 28-OCT-1999; 99US-0161932P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 12.5%; Score 430; DB 3; Length 304;  
 Best Local Similarity 31.2%; Pred. No. 1.9e-31;  
 Matches 102; Conservative 61; Mismatches 110; Indels 54; Gaps 10;

QY 90 DIVNAKIELPIRKNHIDLLIDVFRKKR---GNRLHPKVALLRLRAASLKLQPNIS--- 143  
 DB 8 DIIRLVB-----FRNTRPSGKQVH-----LSHG--EIKROLCAVSKEI 44

QY 144 ----PVSTAVSQVTVCGDLHGKLDLVLVHLKNGLPSSSNPFYVCGDFVDRGKGLVLT 199  
 DB 45 FLQGNPLLEBAPRICGDHIGQYSDLRLRPFYGGFPPEAN-YLTLGQVYDRGKSLRTI 103

QY 200 LLLLSLYLAFPNVAVPLNQNHEDSVMNARYGFIRESVKYPRNHRILAFIDEVYRWLPL 259  
 DB 104 CILLAYKIKYPRNFPLRGNHESASINRIYGFYDECKRF--NVLWKIFPDCNCLPV 160

QY 260 GSVLNSRYLVHGGES-DSTSLDLKSIDRGKYVSIARPLDGRPLDXTENQQLFDIMW 318  
 DB 161 AALIDRLICWGGISPELKSIDQIRNTAR-----PMQIPSSGLVCDLWL 205

QY 319 SDPOATWGCVPNTLUGAGVWFGPDVTDNFLQHRRLSYVIRSHCKRNGHIEFMHDKIITI 378  
 DB 206 SDPSGDVGMKAND-RGVETTRGADKVAFLFKHMDLICRAHQVVDGIEFFAERQVTV 264

QY 379 FSASNYTVAIGSNKAYIRLNNQMLMPH 405  
 DB 265 FSAPNYCGEFDNAGAMMSIDSLMCSF 291

RESULT 37  
 ADT57338 standard; protein; 325 AA.  
 ID ADT57338;  
 AC ADT57338;  
 XX 13-JAN-2005 (first entry)  
 XX Plant polypeptide, SEQ ID 7415.  
 DS Plant polypeptide, SEQ ID 7415.  
 KW Plant; transgenic; cold tolerance; growth rate; drought tolerance;  
 KW disease resistance; galactomannan production; plant growth regulator;  
 KW heat tolerance; herbicide tolerance; lignin production;  
 KW extreme osmotic condition tolerance; pathogens resistance;  
 KW pest resistance; yield improvement; seed oil yield; seed protein yield.  
 XX Viridiplantae.



XX	US2004216190-A1.
PN	.
PD	28-OCT-2004.
PF	18-DEC-2003; 2003US-00739930.
PR	28-APR-2003; 2003US-00424599.
PS	28-APR-2003; 2003US-00425115.
PT	(KOVA/) KOVALIC D K.
XX	Kovalic DK;
XX	WPI; 2004-757369/74.
XX	New recombinant DNA constructs useful in the field of biochemistry and genetics, and in particular for producing transgenic plants with improved biological characteristics.
XX	Claim 2; SEQ ID NO 7415; 14pp; English.
CC	The invention relates a recombinant DNA construct comprising a polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO: 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean, Arabidopsis, wheat and rape but the specification does not indicate which sequences is derived from which organism. Also included is a method of producing a plant having an improved property, comprising transforming a plant with a recombinant DNA construct comprising a promoter region functional in a plant cell operably joined to a polynucleotide encoding a polypeptide associated with the property, and growing the transformed plant. The property is selected from improving plant cold tolerance, for manipulating growth rate in plant cells by modification of the cell cycle pathway, for improving plant drought tolerance, for providing increased resistance to plant disease, for galactomanan production, for production of plant growth regulators, for improving plant heat tolerance, for improving plant tolerance to herbicides, for increasing the rate of homologous recombination in plants, for lignin production, for improving plant tolerance to extreme osmotic conditions, for improving plant tolerance to pathogens or pests, for yield improvement by modification of photosynthesis, for modifying seed oil yield and/or content, for modifying seed protein yield and/or content, for yield improvement by modification of carbohydrate, nitrogen or phosphorus use and/or uptake and for yield improvement by providing improved plant growth and development under at least one stress condition. The polynucleotide may also encode a plant transcription factor. The methods and compositions of the present invention are useful in the field of biochemistry and genetics, in particular for producing transgenic plants with improved biological characteristics such as increased yield, improved nitrogen flow, increasing plant tolerance to cold or heat, improving plant tolerance to extreme osmotic and drought conditions, and improving plant tolerance to plant pests or pathogens. They can also be used in physical arrays of molecules, plant breeding markers, computer-based storage and analysis systems. The present sequence is one of the 5544 plant protein sequences of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/Sequence.html?DocID=20040216190.
CC	Sequence 325 AA;
QY	Query Match 12.5%; Score 429.5; DB 8; Length 325; Best Local Similarity 31.6%; Pred. No. 2.4e-11; Matches 99; Conservative 67; Mismatches 118; Indels 29; Gaps 6
DB	94 AKIEPIRKNIHDLIDVFRKRGGRNBLHPKYVALILREAKSLIKOLNPISPVSTAVSQGV 153 14 AAVDEVRR-----LYBGRGCGQGVOVMSEAIRIQLCEBAKQVLLSQPNL-----LRTHAPV 64
QY	154 TVCGDLHGKLDLLLVTLAHKNGLPSSNPVFNQDPVDKRGKGLEALLLSLTLPANAV 213 65 KTCGGIHGGFVVLLRLFDGSGYPPTGT-YIFLADVDYDRGQSLEETCLLAAYLKPYDNI 1233

QY	214	FLNNGNHDSSVNNAAVYGFRLREVSQSPRNHKKILAFIDSVYMWLPFGSVLANSVLVHGG	273
DB	124	YLNKGNHSDAKINRYVGYDECKRRF---NVALMKIPCDCFCNCPALAIIDDKILCMHGG	180
QY	274	FS-DSTSIDLTKSIDRGKVVSLIRPLDTGBPLDKTERWQOIPIDMWSDPQATWGCVPNTL	332
DB	181	LSPELTSLDQIKDIR-----PTETPDVGLLDDLLMSDSPHTBEGWBSQ	225
QY	333	RGAGCWFGPDYTDNPLQNNRLSYVIRSHCEKNGHGFPHNDKILITTPSASNTYAIGSNKG	392
DB	226	RGVSCFTGADKLYEFLTKRNDLIDVGRAGHVVEDGYEFERRLVTLTFSPAPNYCGEFDNAG	285
QY	393	AYIRLNQNLMPHF 405	
DB	286	ALLSIDBSLMCSF 298	
RESULT 38			
ADTS9253			
ID	ADTS9253	standard; protein; 326 AA.	
XX	ADTS9253;		
XX	13-JAN-2005	(first entry)	
DE	Plant polypeptide, SEQ ID 9330.		
XX	Plant; transgenic; cold tolerance; growth rate; drought tolerance;		
KW	disease resistance; galactomannan production; plant growth regulator;		
KW	heat tolerance; herbicide tolerance; lignin production;		
KW	extreme osmotic condition tolerance; pathogens resistance;		
KW	pest resistance; yield improvement; seed oil yield; seed protein yield.		
OS	Viridiplantae.		
XX	US2004216190-A1.		
PN	28-DEC-2003; 2003US-00739930.		
XX	18-DEC-2003; 2003US-00739930.		
PF	28-APR-2003; 2003US-00424599.		
PR	28-APR-2003; 2003US-00425115.		
XX	(KOVA/) KOVALIC D K.		
PA	Kovalic DK;		
XX	Kovalic DK;		
PI	WPI; 2004-757369/74.		
DR	New recombinant DNA constructs useful in the field of biochemistry and		
XX	genetics, and in particular for producing transgenic plants with improved		
PT	biological characteristics.		
PS	Claim 2; SEQ ID NO 9330; 14pp; English.		
CC	The invention relates a recombinant DNA construct comprising a		
CC	polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:		
CC	1-5544) and encoding a polypeptide with any of 5544 amino acid sequences		
CC	(SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,		
CC	Arabidopsis, wheat and rape but the specification does not indicate which		
CC	sequences is derived from which organism. Also included is a method of		
CC	producing a plant having an improved property, comprising transforming a		
CC	plant with a recombinant DNA construct comprising a promoter region		
CC	functional in a plant cell operably joined to a polynucleotide encoding a		
CC	polypeptide associated with the property, and growing the transformed		
CC	plant. The property is selected from improving plant cold tolerance, for		
CC	manipulating growth rate in plant cells by modification of the cell cycle		
CC	pathway, for improving plant drought tolerance, for providing increased		
CC	resistance to plant disease, for galactomannan production, for production		
CC	of plant growth regulators, for improving plant heat tolerance, for		
CC	improving plant tolerance to herbicides, for increasing the rate of		



QY 382 SNYYAIGSNKGAVYIRLNQIMPH 405  
 DB 281 PNYCGEFDNAGALMSIDNSLVCSP 304

RESULT 40

ID ADY06935 standard; protein; 338 AA.

ADY06935;

21-APR-2005 (first entry)

DB Plant full length insert polypeptide seqid 62750.

XX plant protectant; plant growth regulant; gene therapy; plant;  
 XX recombinant DNA construct; physical array; plant breeding marker;  
 KM cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;  
 KM extreme osmotic condition; pathogen tolerance; pest tolerance;  
 KM growth rate; cell cycle pathway; disease resistance;  
 KM galactomannan production; lignin production; plant growth regulator;  
 KM yield; plant growth; plant development; seed oil; protein yield;  
 KM protein content.

XX Unidentified.

XX US2004034888-A1.

PD 19-FEB-2004.

28-APR-2003; 2003US-00425114.

06-MAY-1999; 99US-00304517.

05-NOV-2001; 2001US-00985678.

XX (LIU/) LIU J.

PA (KOVN/) KOVALIC D K.

PA (SCRE/) SCREEN S E.

PA (TABN/) TABASKA J E.

PA (CAOY/) CAO Y.

PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;

XX WPI; 2004-180133/17.

PT New recombinant DNA construct, useful for improving plant tolerance to  
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or  
 PT pests, for conferring increased resistance to plant disease, or for  
 PT improving yield.

PS Claim 1; SEQ ID NO 62750; 15pp; English.

XX The invention describes a recombinant DNA construct comprising a  
 CC polynucleotide consisting of a sequence encoding an amino acid sequence  
 CC available in electronic form from the US patent office at  
 CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide  
 CC of the invention are also useful in physical arrays of molecules and as  
 CC plant breeding markers. The recombinant DNA construct is useful for  
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme  
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in  
 CC plant cells by modification of the cell cycle pathway, for conferring  
 CC increased resistance to plant disease, for producing galactomannan,  
 CC lignin or plant growth regulators, for increasing the rate of homologous  
 CC recombination in plants, for improving yield by modification of  
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake  
 CC or by providing improved plant growth and development under at least one  
 CC stress condition or for modifying seed oil or protein yield and/or  
 CC content. This is the amino acid sequence of a plant full length insert  
 CC polypeptide that can be used in the recombinant DNA construct of the  
 CC invention.

XX Sequence 338 AA;

Query Match 12.5%; Score 429; DB 8; Length 338;  
 Best Local Similarity 31.8%; Pred. No. 2.9e-31;  
 Matches 103; Conservative 60; Mismatches 129; Indels 32; Gaps 7;

QY 94 AKIEIPR-----KNHIDLLIDVFRKKGRRLHPKVALL---LREAAKSLKQLPNI 142  
 DB 1 AGISLPSRSFPGEMDEAAVDDILIRLLRARGRT-PRNAQYDABIRLCAAAKQVFLS 59  
 QY 143 SPVSTAVSOQVWVCGDLHGKLDLIVLHKNGLPSSSNPYFENGDFVDRGKRLGLVLL 202  
 DB 60 QPVLELRLAPIKICGVHGYSDILRLFEVGYPPRAN-YLRFSDYVDKSGSIFRICLL 118  
 QY 203 LSLVLAFPNAVFLNRGNHEDSVNNARYGFIREVESKYPRNHKRIILAFIDEVYRMLPGSV 262  
 DB 119 LAKIKYPPNPFLLRGNHRCASINRIYGFPEDECKRRF--NVRIMKIFTECFNCLPVAAL 175  
 QY 263 LNSRVLIYHGGPS-DSTSLDLTKSIDRGKYSILRPLDGBPLDKTEMQIFDIWMSDP 321  
 DB 176 IDDKIFCMHGGUSPDIKSMQIRNIPR-----PVDVVDGLCLDLMSDP 220  
 QY 322 QATMGCVPTWLRGAGVGFPGPDVTDNFLQRHRLSYVIRSHCKPNGHAFMHDKITITFSA 381  
 DB 221 DKEIDKMGENDRGVSYTFGADVAAEFLQKHLDLIGRAHQVVEDGYEFPKQQLVITIFSA 280  
 QY 382 SNYYAIGSNKGAVYIRLNQIMPH 405  
 DB 281 PNYCGEFDNAGALMSIDNSLVCSP 304

Search completed: January 20, 2006, 19:48:57  
 Job time : 139 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 20, 2006, 19:45:57 ; Search time 46 Seconds  
(without alignments)  
1188.014 Million cell updates/sec

Title: US-09-463-733-1

Perfect score: 3442  
Sequence: 1 MDENALRALAFIQKWRHQ.....VEHDIDPTDESKVIDPKKS 661

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/ptodata/1/1aa/6.COMB.pep: \*  
2: /cgn2\_6/ptodata/1/1aa/6.COMB.pep: \*  
3: /cgn2\_6/ptodata/1/1aa/H.COMB.pep: \*  
4: /cgn2\_6/ptodata/1/1aa/PCTUS.COMB.pep: \*  
5: /cgn2\_6/ptodata/1/1aa/RB.COMB.pep: \*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	570	16.6	499	2	US-09-949-016-7370 Sequence 7370, App
2	543	15.8	494	2	US-09-517-779-2 Sequence 2, Appli
3	521.5	15.2	494	2	US-09-270-767-43457 Sequence 43457, A
4	431.5	12.5	445	2	US-09-248-796A-18327 Sequence 18327, A
5	418.5	12.2	312	2	US-09-538-092-256 Sequence 256, App
6	412	12.0	319	2	US-09-248-796A-18323 Sequence 18323, A
7	405.5	11.8	323	2	US-09-538-092-1093 Sequence 1093, App
8	403.5	11.7	330	2	US-09-167-206-6 Sequence 6, Appli
9	403.5	11.7	330	2	US-09-538-092-873 Sequence 873, App
10	403	11.7	524	2	US-09-744-016A-33 Sequence 33, Appli
11	403	11.7	692	2	US-09-487-558B-274 Sequence 274, App
12	402.5	11.7	374	2	US-09-949-016-7363 Sequence 7363, App
13	402	11.7	327	2	US-09-538-092-1095 Sequence 1095, App
14	402	11.7	327	2	US-09-949-016-6266 Sequence 6266, App
15	402	11.7	356	2	US-09-949-016-8084 Sequence 8084, App
16	402	11.7	502	2	US-09-949-016-6642 Sequence 6642, App
17	402	11.7	509	2	US-09-744-016A-12 Sequence 12, Appli
18	402	11.7	514	2	US-09-744-016A-30 Sequence 30, Appli
19	402	11.7	535	2	US-09-949-016-11179 Sequence 11179, A
20	395.5	11.5	330	2	US-09-190-976B-8 Sequence 8, Appli
21	395.5	11.5	710	2	US-09-487-558B-276 Sequence 276, App
22	391.5	11.4	502	2	US-09-949-016-10218 Sequence 10218, A
23	391.5	11.4	510	2	US-09-744-016A-6 Sequence 6, Appli
24	391.5	11.4	521	2	US-09-538-092-1283 Sequence 1283, App
25	391.5	11.4	523	2	US-09-744-016A-21 Sequence 21, Appli
26	391.5	11.4	533	2	US-09-744-016A-18 Sequence 18, Appli
27	391	11.4	309	2	US-09-538-092-902 Sequence 902, App

28	391	11.4	309	2	US-09-949-016-6461 Sequence 6461, App
29	391	11.4	311	2	US-09-949-016-11335 Sequence 11335, A
30	386	11.2	341	2	US-09-949-016-8780 Sequence 8780, App
31	384	11.2	355	2	US-09-949-016-11414 Sequence 11414, A
32	383.5	11.1	312	2	US-09-949-016-9733 Sequence 9733, App
33	380	11.0	487	1	US-08-452-722-7 Sequence 7, Appli
34	380	11.0	487	1	US-08-404-731A-7 Sequence 7, Appli
35	380	11.0	487	1	US-08-344-227-7 Sequence 7, Appli
36	380	11.0	487	1	US-08-503-226B-7 Sequence 7, Appli
37	380	11.0	487	2	US-08-721-458B-7 Sequence 7, Appli
38	380	11.0	528	2	US-09-744-016A-9 Sequence 9, Appli
39	380	11.0	535	2	US-09-949-016-11683 Sequence 11683, A
40	380	11.0	535	2	US-09-744-016A-24 Sequence 24, Appli
41	379	11.0	309	2	US-09-538-092-854 Sequence 854, App
42	378	11.0	377	2	US-09-487-558B-270 Sequence 270, App
43	375	10.9	260	2	US-09-248-796A-18344 Sequence 18344, A
44	373.5	10.9	359	2	US-09-487-558B-268 Sequence 268, App
45	372	10.8	306	2	US-09-828-302-13 Sequence 13, Appli

## ALIGNMENTS

```
RESULT 1
US-09-949-016-7370
; Sequence 7370, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTNER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 7370
; LENGTH: 499
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7370
Query Match 16.6%; Score 570; DB 2; Length 499;
Best Local Similarity 33.7%; Pred. No. 2,7e-46;
Matches 143; Conservative 71; Mismatches 142; Indels 68; Gaps 16;
11 FICKWYRQAREREMRCQWQIFQNLVYASDDQALVKFENDLTKHMPQAGRNQYQ 70
95 YIGYVRRRAASNAALAK-----FR-----AALAD-----YETVVKVQPHDQAKKQY 137
71 -----GSAHV-SVLDDKD-----DLVERGQ-DIVAKIELPIRKNHIDL 107
138 ECKVIKQAFERAIAGDEKRSVVDLSIESWTITEDYSGPLDEBKQVITSMKE----- 193
108 LIDVFKKRGKGNRLHPTVALILREAKSLKQLPNISPVSTANSQOVTVCGDLHGKLDL 167
194 LMQWYDQK-KLHRCAYQIOLVQVKEVSKSTLIVETLLKRTIKTVCGDTHGQFPDYL 251
168 VVLAHKRGGLSSNPYVNGDFVDRGKGLVLLLSLVYAPNVAFLRGNHEDSVMA 227
252 NIEELNGLSSEINPIYFNEDFVDRGSPSVVILITGFKLLVPHDHLRGNHETNMNO 311
228 RYGFIREVESKYPNRNKRILAFIDEVYRWLPLGVSVANSVTLVHGGSF--SDSTSLDLIR 285
312 IYGFEGEVAKVY---AQWYELPSVFEWLPPLAQICNGVLLIMHGGLFSEBDGTVLDDIRK 368
286 IDRGKTVSLRPLLTGDEPLDKTEWQQTIDIMMSDPQAMGCVPTNLRGAGVAFGPDVTD 345
```

```
Db 369 IERNR-----QPP--DSGPM-----CDLMSDQPOQG-RSISKRVSCQFSGPDYK 412
Qy 346 NPLQRLRLSVIRSHCKPENGHEFMDNKLITTFASNTYALGSKGAYIRL-NNQJMPH 404
Db 413 AFLERNLNYIIRSHVKAGYEVAGHRCVTFSPAPNYCDQGNKASYYIHLQGSILRQ 472
Qy 405 FVOY 408
Db 473 FHOQ 476

RESULT 2
US-09-517-779-2
; Sequence 2, Application US/09517779
; Patent No. 6660511
; GENERAL INFORMATION:
; APPLICANT: Luo, Xiang
; APPLICANT: Xu, Xiang
; TITLE OF INVENTION: Cell Cycle Proteins Associated with Rad9, Compositions
; TITLE OF INVENTION: and Methods Of Use
; FILE REFERENCE: A68293/RMS/DAY
; CURRENT APPLICATION NUMBER: US/09/517,779
; CURRENT FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 494
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-517-779-2
```

```
Query Match 15.8%; Score 543; DB 2; Length 494;
Best Local Similarity 33.4%; Pred. No. 1.1e-43;
Matches 142; Conservative 71; Mismatches 142; Indels 70; Gaps 18;

Qy 11 PLOKWRRLQARRRMOQRRCNMOIFONLSEYASBODQALYKFFNDLKKHMQAAGRNOYO 70
Db 90 YIKGYRRAASNNALQK-----FR-----AALRD-----YETVVKYKPHDKAKMKQ 132
Qy 71 -----GSAHV-SVLDDKD---DLVEFPG-DIVAKIELPIKKNHDL 107
Db 133 ECKNIKOKAFERAIAGDEHKRSVVDLIESWTIDREYSGPLDQKVTIISPMKE----- 188
Qy 108 LIDVFKKKKGNRLHPKVALILBAKSLKQLEPNISPVSTASQOYTVCGDLHGKLDLL 167
Db 189 LKQWYDQK--KLHRCAYQILVQKRVLSKSLTIVETTLKTEKITTVC-DTHGQFYDLL 245
Qy 168 VVLHKNGLPSSSNPY-VFNGDFVDRGKRGLEVLILLSLYLAFPNAVFLNRGNHEDSVNN 226
Db 246 NIFELNGLSBETNPYIFINQDFVDRGSPSVIEVILTLFGKRLYDPHFLIRGNHETDNKN 305
Qy 227 ARYGFIRVESKYPNNHKLILAFIDVYRMLPLGSVLSNRVLIHNGF--SDSTSLDLK 284
Db 306 QLYGFEGEYAAKYT--AQWYELFSHFVFWFLAQCIINGVLIIMHGLPSEBDVTLDDIR 362
Qy 285 SIDRGKYVALRPLTDEBPLDKTEWQQLFDIMNSPQATMGCVPTLGLAGATWGPDDYT 344
Db 363 KIERNN-----QPP--DSGPM-----CDLMSDQPOQG-RSISKRVSCQFSGPDYK 406
Qy 345 DNFLQRLRLSVIRSHCKPENGHEFMDNKLITTFASNTYALGSKGAYIRL-NNQJMP 403
Db 407 KAFLENNNDYIIRSHVEVAEGYEVAGHRCVTFSPAPNYCDQGNKASYYIHLQGSILR 466
Qy 404 HFOY 408
Db 467 QFHOQ 471
```

```
RESULT 3
US-09-270-767-43457
; Sequence 43457, Application US/09270767
; Patent No. 6703491
```

```
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43457
; LENGTH: 494
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-43457
```

```
Query Match 15.2%; Score 521.5; DB 2; Length 494;
Best Local Similarity 31.2%; Pred. No. 1.4e-41;
Matches 130; Conservative 72; Mismatches 126; Indels 87; Gaps 14;

Qy 4 NAIRA-ALFIQWYRRHQARRM---QRCNMOIFONLSEYASBODQALYKFFNDLKKH 58
Db 119 SAVKADPAPYLKGYRRAAHSLSGFKQALCDPEFVAKCR-FNDQAKLKFTCKRYK- 176
Qy 59 MQQAAGRNQYGSAAHVSGLDQKDLVEBFGDIVAKTELPKKNHIDLIDVFKKGN 118
Db 177 -MRAFERA-----IADKPEKTLSEYSDMENITIE-----DDY---KGP 212
Qy 119 RLHPRYVALILREAKSLKQLE-----PNISPVSTAV 149
Db 213 QLEDKQVTL-----KFMKELAKXXXXXXXXXXXXXXXXXXXXXXXXXXXXMBAQPLVDITVPD 266
Qy 150 SCQYTVCGDLHGKLDLVLVHLKNGLPSSSNPYVNGSPVDRGKRGLEVLILLSLYLA 209
Db 267 EEKFTICDHIQOQFYDLNMLFEINGLPSBKPYLFPNGDFVDRGSPSVICFTLPFGKLY 326
Qy 210 PNAVLNRGNHEDSVNARVYGFIRVESKYPNNHKLILAFIDVYRMLPLGSVLSNRVLI 269
Db 327 PNHFLAAGNHESIMNOMTGFTGYTAKYT---SAMADIFQVNMPLCHCINOKTIV 383
Qy 270 VHGF--SDSTSLDLKSIDRGKYVSLRPLTDEBPLDKTEWQQLFDIMNSPQATMG 327
Db 384 HNGGLFSTEDVTLDIRIRERN-----CQPP-----EBGLMCELMSDPOQMG- 427
Qy 328 VENTIRGAGWGPVPTDNLQRLRLSVIRSHCKPENGHEFMDNKLITTFASANTY 384
Db 428 LQSGKRGVGIQGPVTEKFCNNMLDYIIRSHVEYDMGYEVAGHNGKCIYFASANTY 484
```

```
RESULT 4
US-09-248-796A-18327
; Sequence 18327, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18327
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (220), (239)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknc
```

US-09-248-796A-18327

```
Query Match      12.5%; Score 431.5; DB 2; Length 445;
Best Local Similarity 33.2%; Pred. No. 6.5e-33;
Matches 104; Conservative 57; Mismatches 105; Indels 47; Gaps 7;

QY 103 NHIDLIDVPRKRGKGNLHPKYVALILREAAKSLKOLPNISPVSTAVSQ----- 151
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 132 DYLMKTKLP--KNKGKLPKSHVPAIIAKYELIKONTMTETSLHSQIDFNNDNDT 189
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 152 -----QVTVCGDLHGKLDLVLVHKNGLPSSSNPYVNGDPVDRGK 193
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 190 NNPHGKRGKSDQIIGKGLTVGDTHGQPYDLNLDPDKGHVQDHIIYKENGVDVDRGS 249
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 194 RGLVVLILLISLYAFPNNAFLNKGHEDSVNARYCFIREVESKPRNHRILAFIDEV 253
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 250 WSCVVALYLVKLIVPKSIFINNGHNTDMNTYGFNDECEPKY--SKKIPEAFNQS 306
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 254 YRWLPISVNSRVLIVHGGF--SDSTSLDKSIDGKYSILRPLDTGEPDLKTEMQ 311
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 307 FGALPLACLINQYVLCNHGGLPCNDKVTLDKDKISINPHSGSQPP--KEGLAV----- 359
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 312 QIPDMSDPOATWGCVPNTLRGAGVFGPDVTNPLQRRRLSVYIRSHCEKPGHEFMH 371
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 360 ---ELMTWDPQEIINGSSPSK--RGIGMQPGDITRFCLSKIKIKIISHVVRANGINEQ 415
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 372 DNKIIITFSASNY 384
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 416 NGRIMTVFSAFNY 428
```

RESULT 5

```
US-09-538-092-256
; Sequence 256, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Glaxo, Ltd
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538, 092
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurabacSeqFormatter Version 0.9
; SEQ ID NO 256
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number YER133W
US-09-538-092-256
```

```
Query Match      12.2%; Score 418.5; DB 2; Length 312;
Best Local Similarity 31.5%; Pred. No. 6.6e-32;
Matches 100; Conservative 63; Mismatches 125; Indels 29; Gaps 9;

QY 103 NHIDLIDVPRKRGKGN--LHPKYVALILREAAK--SLKOLPNISPVSTAVSQVTVCGDL 159
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 10 NIIDRLLEVRSKRGQGVDLLENIRYLCSKARSIFIKQ-----PILBLEAPIKICGI 64
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 160 HGKLDLILVHLKNGLPSSSNPYVNGDPVDRGKGLGVLLILLISLYAFPNNAFLNKG 219
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 65 HGQYDYLRLRFERYGFPPEEN--YLFIDGYDRGQSLETTICLLAYKIKYPENFFILRGN 123
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 220 HEDSVNARYGFIREVSKYPRNHRILAFIDEVYRWLPISVNSRVLIVHGGFS--DST 278
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 124 HECASIRIRIYGFYDCKRRY--NIKLMKTPTD--CFNCLPAAIIDEKIFCMHGGLSFDLN 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
QY 279 SLDLIKSIDRGKYSILRPLDTGEPDLKTEMQIIPDMSDPOATWGCVPNTLRGAGV 338
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 SMEQIRRVNR-----PTDIPVGLICDLMSDDPKDIVGSENDRGVSFT 225
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 339 FGPDPVTNPLQRRRLSVYIRSHCEKPGHEFMHDKIITTFSSNYYAIGSNKAYIRLN 398
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 226 FGPDVNRRFLQKQDMELICRAHQVEDGYEFPSSKQDVLTPSAFNYCGEFPDNGAMMSVD 285
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 399 NQLMPHFVOYIYSAASQT 415
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 286 ESLICSF--QILKPAQKS 301
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 6

```
US-09-248-796A-18323
; Sequence 18323, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196,132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18323
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (16)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unk
US-09-248-796A-18323
```

```
Query Match      12.0%; Score 412; DB 2; Length 319;
Best Local Similarity 31.2%; Pred. No. 2.9e-31;
Matches 95; Conservative 60; Mismatches 125; Indels 24; Gaps 7;
```

```
QY 118 NRIHPKYVALI---LREAAKSLKOLPNISPVSTAVSQVTVCGDLHGKLDLILVHLKNG 174
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 13 NRHXGRVTLTEIRHIFLCTKAREIFIQPILBLEAPIKICGDIHGQYDYLRLRFYVG 72
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 175 LPSSSNPYVNGDPVDRGKGLGVLLILLISLYAFPNNAFLNKGHEDSVNARYGFIIR 234
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 73 FPEEAN--YLFIDGYDRGQSLETTICLLAYKIKYPENFFILRGNHECASINRIYGFYD 131
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 235 VESKYPNHRILAFIDEVYRWLPISVNSRVLIVHGGFS--DSTSLDLIKSIDRGKYS 293
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 132 CKRRP--NIKLMKTPTD--CFNCLPAAIIDEKIFTHGGLSFDLMSWEQIRRVNR----- 183
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 294 ILRPLDTGEPDLKTEMQIIPDMSDPOATWGCVPNTLRGAGVFGPDVTNPLQRRRL 353
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 184 -----PDIIPVGLICDLMSDDPKDITGSENDRGVSFTFGPVSFRLQKDM 233
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 354 SVYIRSHCEKPGHEFMHDKIITTFSSNYYAIGSNKAYIRLNQLMHFVOYIYSAAS 413
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 234 DLICRAHQVEDGYEFPSSKQDVLTPSAFNYCGEFPDNGAMMSVDLSLCSF--QILKPAD 292
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 414 QTKR 417
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 293 KKPR 296
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 7

```
US-09-538-092-1093
; Sequence 1093, Application US/09538092
; Patent No. 6753314
```

```
/ GENERAL INFORMATION:
/ APPLICANT: Glot, Loic
/ APPLICANT: Mansfield, Traci A.
/ TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
/ FILE REFERENCE: 15966-542
/ CURRENT APPLICATION NUMBER: US/09/538,092
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: 60/127,352
/ PRIOR FILING DATE: 1999-04-01
/ PRIOR APPLICATION NUMBER: 60/178,965
/ PRIOR FILING DATE: 2000-02-01
/ NUMBER OF SEQ ID NOS: 1387
/ SOFTWARE: CuraPatSeqFormatter Version 0.9
/ SEQ ID NO 1093
/ LENGTH: 323
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (0)...(0)
/ OTHER INFORMATION: Polypeptide Accession Number P36873
US-09-538-092-1093
```

```
Query Match      11.8%; Score 405.5; DB 2; Length 323;
Best Local Similarity 32.4%; Pred. No. 1.3e-30;
Matches 105; Conservative 58; Mismatches 126; Indels 35; Gaps 11;
```

```
QY 102 KKHIDLLIDVFRKKGNR-----LHPKYV-ALILBAKSLKQLNISPVTAVSQOYV 154
DB 6 KNLDSIIQRLLEVGQSR-PGKNVQLTNEIRGLCLKRSREIFLSQ-----PILILEAPLTK 60
QY 155 VCGDLHGKLDLVLVHKNGLPSSSNPYVFNQDFVDRGKRGLEVLILLSLYAFPNAYFL 214
DB 61 ICGDHGQYVDLRLFEYGGFPPESSN-YFLGIDYDRGQSLTICLLAYKIKYENF 119
QY 215 LNFGEHDSVMNARYGFIREVSKYPRNKRILAFIDEYVRMLPLGSVLSRVLIVHGCF 274
DB 120 LLEGNHCASINRIYGFYDECKRRY--NITLMTFTD-CFNCPLAIAYDEKIFCGGGL 176
QY 275 S-DSTSLDIKSIDRGKYVSIILRPPLTDSPLDKTEWQOIFDIWMSDP-QATMGCVPTNL 332
DB 177 SPLQSMEOIRRLMR-----PTVPPDGLLCDLWMDPDKDVLGMSGEND- 220
QY 333 RGAGVGFDPVTNFIQRLRLSYVIRSHCKPGHGFMDNKIITIFSASNYVAISNGK 392
DB 221 RGVSTFGAEVAKFLKHDLIDICRAHQYVEDGYEFPAKQVLVTLFSAPNTGSEFDNAG 280
QY 393 AYIRLNNQIMPHFVQYISAASQTK 416
DB 281 AMMSVDETLWCSP-QILKPAKCK 303
```

```
RESULT 8
US-09-167-206-6
/ Sequence 6, Application US/09167206A
/ Patent No. 6476193
/ GENERAL INFORMATION:
/ APPLICANT: Nandabalan, Krishnan
/ APPLICANT: Schulz, Vincent P.
/ APPLICANT: Yang, MeiJa
/ TITLE OF INVENTION: NIK1 PROTEIN AND NIK1 PROTEIN COMPLEXES
/ FILE REFERENCE: 15966-521 NIK1 protein complexes
/ CURRENT APPLICATION NUMBER: US/09/167,206A
/ PRIOR FILING DATE: 1998-10-06
/ NUMBER OF SEQ ID NOS: 26
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 6
/ LENGTH: 330
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-167-206-6
```

```
Query Match      11.7%; Score 403.5; DB 2; Length 330;
```

```
Best Local Similarity 31.7%; Pred. No. 2.1e-30;
Matches 101; Conservative 60; Mismatches 133; Indels 25; Gaps 8;
QY 102 KKHIDLLIDVFRKKGNRLHPKYVALI---LREAKSLKQLNISPVTAVSQOYVCGD 158
DB 6 KNLDSIIQRLLEVGQSR-PGKNVQLTNEIRGLCLKRSREIFLSQPIILILEAPLTKICGD 64
QY 159 LHGKLDLVLVHKNGLPSSSNPYVFNQDFVDRGKRGLEVLILLSLYAFPNAYFLNAG 218
DB 65 IHGQYVDLRLFEYGGFPPESSN-YFLGIDYDRGQSLTICLLAYKIKYENFLLNG 123
QY 219 NHEDSVNARVGFIREVSKYPRNKRILAFIDEYVRMLPLGSVLSRVLIVHGGS-DS 277
DB 124 NHCASINRIYGFYDECKRRY--NITLMTFTD-CFNCPLAIAYDEKIFCGGGLSPDL 180
QY 278 TSLDLIKSIDRGKYVSIILRPPLTDSPLDKTEWQOIFDIWMSDPQATMGCVPTNLRGAGV 337
DB 181 QSMEOIRRLMR-----PTVPPDGLLCDLWMDPDKDVLGMSGENDRGVSF 225
QY 338 WFGPVTNFIQRLRLSYVIRSHCKPGHGFMDNKIITIFSASNYVAISNGKAYIRL 397
DB 226 TFGAEVAKFLKHDLIDICRAHQYVEDGYEFPAKQVLVTLFSAPNTGSEFDNAGMSV 285
QY 398 NNQIMPHFVQYISAASQTK 416
DB 286 DETLWCSP-QILKPADKCK 303
```

```
RESULT 9
US-09-538-092-873
/ Sequence 873, Application US/09538092
/ Patent No. 6753314
/ GENERAL INFORMATION:
/ APPLICANT: Glot, Loic
/ APPLICANT: Mansfield, Traci A.
/ TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
/ FILE REFERENCE: 15966-542
/ CURRENT APPLICATION NUMBER: US/09/538,092
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: 60/127,352
/ PRIOR FILING DATE: 1999-04-01
/ PRIOR APPLICATION NUMBER: 60/178,965
/ PRIOR FILING DATE: 2000-02-01
/ NUMBER OF SEQ ID NOS: 1387
/ SOFTWARE: CuraPatSeqFormatter Version 0.9
/ SEQ ID NO 873
/ LENGTH: 330
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (0)...(0)
/ OTHER INFORMATION: Polypeptide Accession Number P08129
US-09-538-092-873
```

```
Query Match      11.7%; Score 403.5; DB 2; Length 330;
```

```
Best Local Similarity 31.7%; Pred. No. 2.1e-30;
```

```
Matches 101; Conservative 60; Mismatches 133; Indels 25; Gaps 8;
```

```
QY 102 KKHIDLLIDVFRKKGNRLHPKYVALI---LREAKSLKQLNISPVTAVSQOYVCGD 158
DB 6 KNLDSIIQRLLEVGQSR-PGKNVQLTNEIRGLCLKRSREIFLSQPIILILEAPLTKICGD 64
QY 159 LHGKLDLVLVHKNGLPSSSNPYVFNQDFVDRGKRGLEVLILLSLYAFPNAYFLNAG 218
DB 65 IHGQYVDLRLFEYGGFPPESSN-YFLGIDYDRGQSLTICLLAYKIKYENFLLNG 123
QY 219 NHEDSVNARVGFIREVSKYPRNKRILAFIDEYVRMLPLGSVLSRVLIVHGGS-DS 277
DB 124 NHCASINRIYGFYDECKRRY--NITLMTFTD-CFNCPLAIAYDEKIFCGGGLSPDL 180
QY 278 TSLDLIKSIDRGKYVSIILRPPLTDSPLDKTEWQOIFDIWMSDPQATMGCVPTNLRGAGV 337
```



Db 181 QSMQIRRIIR-----PTDVPQGLCDLLMSDPDXDVGQMGENDRGVSF 225  
 Qy 338 WREGPDYDNLQORHLSYVIRSHCKPNGHGFHNDKIIITPSASNYAIGSKKAYIKL 337  
 Db 226 TFGAEVVAKFLHGDLDLCRAHQVVDGYEFPAKQOLVTLFSAIPYCGFPAAGAMSV 285  
 Qy 398 NNQAMPFVOYISAASQTK 416  
 Db 286 DETLMCSF-QILKPADKXK 303

RESULT 10  
 US-09-744-016A-33  
 ; Sequence 33, Application US/09744016A  
 ; Patent No. 6875581  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dr. Voelkel, Helge  
 ; TITLE OF INVENTION: Method for screening of modulators of calcineurin  
 ; TITLE OF INVENTION: Method for screening of modulators of calcineurin  
 ; FILE REFERENCE: A34157PCT  
 ; CURRENT APPLICATION NUMBER: US/09/744,016A  
 ; CURRENT FILING DATE: 2001-10-03  
 ; PRIOR APPLICATION NUMBER: EP98113876  
 ; PRIOR FILING DATE: 1998-07-22  
 ; NUMBER OF SEQ ID NOS: 35  
 ; SOFTWARE: Patentin Ver. 2.1  
 ; SEQ ID NO 33  
 ; LENGTH: 524  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-744-016A-33

Query Match 11.7%; Score 403; DB 2; Length 524;  
 Best Local Similarity 25.3%; Pred. No. 5.1e-30;  
 Matches 136; Conservative 78; Mismatches 202; Indels 122; Gaps 16;

Qy 93 NAKIEPIRKNIHDLIDVRKKRGNRIHPKXVALILREAKSLKQLPNISPVSTAVSQ 152  
 Db 46 NGEPKVDLKNH-----VKEG-RLBEEVALKINDGAILRQ---EKTMEVDAP 92  
 Qy 153 VIVCGDLHGKLDLVLVHKNGLPSSSNPYVFNQDVRGKGLLEVLLLSLYLAPNA 212  
 Db 93 IYVCGIHQGFPLMLFVGVGSPSNTR-YLFLGDYVDRGYFSECVLYIMSLKIHPT 151  
 Qy 213 VFLRGNHEDSVMMARVGFIRVYSKYPRNKRILAFIDBYRMLPLGVSINRVLIVG 272  
 Db 152 LFLRGNHCRHLTDYFTFKQCRICKY--SEQVYDACMETFDCLPLAALLNQFLCVHG 208  
 Qy 273 GRS-DTSLDLYSIDRGKVSILRPLDGBPLDKTBMQOIPDIMSDPQATMG----- 326  
 Db 209 GMSPEITSLDDIRKLR-----FTERP-----AFGPVCDLLMSDPSSEDDYGNKXTL 253  
 Qy 327 --CVPTLRGAGVFGPDYDNLQORHLSYVIRSHCKPNGHGFHNDK-----ITPT 378  
 Db 254 EHTHTHTVAGCSFYGYPAVCEFLQNNLLSTIRAHADAGYRMRKSGOATGPPSLITI 313  
 Qy 379 FSAASYAIGSNKAVIRLNNQMLPHFVOYISAASQTKLSFKQRMGIYESSALKEIAVR 438  
 Db 314 FSAIPYLDVYNNCAAVLKXENNVM----- 337  
 Qy 439 KMDHRELBDEFKPYKPOSGYISISHWCKVMENVTKGLPMLDLKAPGDSOKNY 498  
 Db 338 -----NIRQFCSPPHY-----WLPNFVDVTWLSLP-----VGEKVEMLV 374  
 Qy 499 NRTLDLDTVILEAADG--MSYMDALYANK-----ASLVAFNIIDADNSGRTLDPEPT 553  
 Db 375 NVANICSDDELIDDDABEGSTTVRKELIRKIRAIKMARVFSITLQGBSSVTLTGLRP 434  
 Qy 554 AIDLVAHMPGAVSKAEMLEKCRMDLNGDKVDLNEFLAFLSLDLHREKQODENIR 611  
 Db 435 TGTLPGLVSGKQTE-----TAKQEAABERRAIGFSLQHKIRSFEBAR 480

RESULT 11  
 US-09-487-558B-274  
 ; Sequence 274, Application US/09487558B  
 ; Patent No. 6949356  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Busby, Robert  
 ; APPLICANT: Cali, Brian  
 ; APPLICANT: Hecht, Peter  
 ; APPLICANT: Holtzman, Doug  
 ; APPLICANT: Madden, Kevin  
 ; APPLICANT: Milne, Todd  
 ; APPLICANT: No. 6949356man, Thea  
 ; APPLICANT: Royer, John  
 ; APPLICANT: Salama, Sofie  
 ; APPLICANT: Sherman, Amir  
 ; APPLICANT: Silva, Jeff  
 ; APPLICANT: Summers, Eric  
 ; TITLE OF INVENTION: Method for Improving Secondary Metabolite Production in Fungi  
 ; FILE REFERENCE: 109272.130  
 ; CURRENT APPLICATION NUMBER: US/09/487,558B  
 ; CURRENT FILING DATE: 2000-01-19  
 ; PRIOR APPLICATION NUMBER: US 60/487,558  
 ; PRIOR FILING DATE: 1999-10-20  
 ; NUMBER OF SEQ ID NOS: 446  
 ; SOFTWARE: Patentin version 3.0  
 ; SEQ ID NO 274  
 ; LENGTH: 692  
 ; TYPE: PRT  
 ; ORGANISM: Saccharomyces cerevisiae  
 US-09-487-558B-274

Query Match 11.7%; Score 403; DB 2; Length 692;  
 Best Local Similarity 28.5%; Pred. No. 8.1e-30;  
 Matches 120; Conservative 71; Mismatches 166; Indels 64; Gaps 12;

Qy 38 EYASEDDQELKFFRDLIKHPQAAAGKQYQGSAAHVSLDQXDLVEFGDIYVA--- 94  
 Db 299 KHIISDDDIENSQSLSNHASMENVNDKXN-----NITSKQDPNEFNDIMOSGN 350  
 Qy 95 -----KIELEPIRKNIH-IDLLIDV-FPKRGNRLHPK-----YVALILREAKSLKQLPNT 142  
 Db 351 KNAKPKFKPIIDETIQCLDAGYAAKTRVCLKNNEIILQICRAKIFLSQPSLEL 410  
 Qy 143 SPVSTAVSQVTVCGDLHGKLDLVLVHKNGLPSSSNPYVFNQDVRGKGLLEVLLL 202  
 Db 411 SP-----PVKIVGVHGYQDLDLRLFTKCGFPSSN-YLFLGDYVDRGKQSLERTILL 462  
 Qy 203 LSLYLAFPAVPLRGNHEDSVMMARVGFIRVYSKYPRNKRILAFIDBYRMLPLGSV 262  
 Db 463 FCYKIKYENPFLLRGNHBCANVTRVGYFDECKKRC-NIKIMKTFID-TFVTLPLAAI 519  
 Qy 263 LNSRVLYVHGSDSTSLDLYSIDRGKVSILRPLDGBPLDKTBMQOIPDIMSDPQ 322  
 Db 520 VAGKIFCHGGLS-----PVLNSMDEIRRV-----RPLDVPDGLINDLMSDPT 565  
 Qy 322 ATWGCVPNTLRGAGVFGPDYDNLQORHLSYVIRSHCKPNGHGFHNDKIIITPSAS 382  
 Db 566 DDPNEMEDBERGVSCYNNVAILNKFLPKRGFPLVCAHVVVDGYEFPRDRSLVTFSPAR 625  
 Qy 383 NYTAAIGSNKAVIRLNNQMLPHFVOYISAASQTKLSFKQRMGIYESSALKEIAVR 442  
 Db 626 NYCGERDMNGAVMSVBGLCSF-----ELDPLDLSAALKQVMKKGRGB 669  
 Qy 443 R 443  
 Db 670 R 670

RESULT 12  
 US-09-949-016-7363  
 ; Sequence 7363, Application US/09949016  
 ; Patent No. 6812339







```

; OPERATING SYSTEM: MS DOS
; SOFTWARE: Microsoft Word version 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/190,976B
; FILING DATE: 12-NO. 6815187-1998
; CLASSIFICATION: Unknown
; ATTORNEY/AGENT INFORMATION:
;   NAME: David Praebker, Esq.
;   REGISTRATION NUMBER: 29,693
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (978) 525-3794
; INFORMATION FOR SEQ ID NO: 8:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 330 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-190-976B-8

Query Match      11.5%; Score 395.5; DB 2; Length 330;
Best Local Similarity 31.3%; Pred. No. 1.3e-29;
Matches 100; Conservative 60; Mismatches 134; Indels 25; Gaps 8;

QY      102 KKHIDLLIDVFRKKGRNRLHPKYVALI---LREAAKSLKQLPNISPVSTAVSQQVTVCGD 158
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB      6 KAMLDSTIGRLHLEVGQSR-PGKNVQLTENEIRGLCLKSRREIFLSQPIILILBAPLKICGD 64

QY      159 LHGKLDLLVVLHKNGLPSSSNPYFVNGDVRGKGLGVLLLLSLYLAFPNAAVFLRNGH 218
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB      65 IHQYQYDLMLFYYGGPPPSN-YLFLGDYDREKQSLFTICLLAKIKYPPNFFLRG 123

QY      219 NHEDSVNNAVGFIREVESKYPNNHKLILAFIDEVYRWLPGLSVLNSRVLIHGGFS-DS 277
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB      124 NHHCASINRIYGYDECKRY--NIKMKFTFD-CFNCPLIAIVDEKIFCGHGLSPDL 180

QY      278 TSLDLKSIDRGKYVSLRPPLTDBEPDTEMQOIFDIMSPPQATMGCVPTLRGAV 337
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB      181 QSMEOIRIRIR-----PTDVPDQGLLDLIMSDPDKQVQCMGRDGVSR 225

QY      338 WFGPDVTDNFLOHRRLSYVIRSHCKPGHGFPHNDKILITFSASNYAIGSKGAYIRL 397
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB      226 TFGAEVVAKFLHGHDLICRAHQVVEDGYEFPKQQLVTLFSAPNCGEFDNAGAMSV 285

QY      398 NNQIMPHFYQYISAASQTK 416
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB      286 DETLMCSF-QILKRPADKMK 303

RESULT 21
US-09-487-558B-276
; Sequence 276, Application US/09487558B
; Patent No. 6949356
; GENERAL INFORMATION:
; APPLICANT: Bueby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. 6949356man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.130
; CURRENT APPLICATION NUMBER: US/09/487,558B
; CURRENT FILING DATE: 2000-01-19
```

```

; PRIOR APPLICATION NUMBER: US 60/487,558
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 446
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 276
; LENGTH: 710
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-487-558B-276

Query Match      11.5%; Score 395.5; DB 2; Length 710;
Best Local Similarity 28.3%; Pred. No. 4.5e-29;
Matches 114; Conservative 72; Mismatches 162; Indels 55; Gaps 11;

QY      60 PQAAERKQYQSA---HVSUDDKDDLYE-----EFGDIYNAKIEL---PIRKNI 105
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB      339 PRKGQRSSSSSSSQRIYVSPSPGDFVHGSCADGDSRNTTWEMRKKEKVPSPVDI 398

QY      106 D-----LLIDVPRKKGRNRLHPKYVALILREAAKSLKQLPNISPVSTAVSQQVTVCGDLH 160
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB      399 DEITQRLDAGYAAKTKNVCLNSHIT--QICHKAREFLAQPALLESPSVKIVGDVH 456

QY      161 GKLDLLVVLHKNGLPSSSNPYFVNGDVRGKGLGVLLLLSLYLAFPNAAVFLRNGH 220
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB      457 GQYADLLRLFTKCGFPPMAN-YLFLGDYDREKQSLFTICLLAKIKYPPNFFLRG 515

QY      221 EDSVNAVGFIREVESKYPNNHKLILAFIDEVYRWLPGLSVLNSRVLIHGGFSDSTL 280
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB      516 ECANVTRVYGFYDECKRC--NIKMKTFVD--TFNTLPLAIVTGKIFCVHGGLS----- 567

QY      281 DLKSIDRGKYVSLRPPLTDBEPDTEMQOIFDIMSPPQATMGCVPTLRGAV 340
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB      568 PVLNSWDEIRHVS-----RPTDVPDQGLINDLIMSDPTSSNMEWENEGVSPCYN 618

QY      341 PVDTDNPLORHRLSYVIRSHCKPGHGFPHNDKILITFSASNYAIGSKGAYIRLNG 400
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB      619 KVAINFELKFGFDYCRAMVYVEDGYEFPNDSLVTVBSAPYCGEFDWGVAVTVS 678

QY      401 LMPHFVQYISAASQTKRLSFKQMGIVSSALKELAVRMEDHR 443
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB      679 LILCSF-----BLDPLDSTALAKQVMKKGRQR 705

RESULT 22
US-09-949-016-10218
; Sequence 10218, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10218
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10218

Query Match      11.4%; Score 391.5; DB 2; Length 502;
Best Local Similarity 31.1%; Pred. No. 6.2e-29;
Matches 101; Conservative 58; Mismatches 119; Indels 47; Gaps 10;

QY      102 KKHIDLLIDVFRKKGRNRLHPKYVALILREAAKSLKQLPNISPVSTAVSQQVTVCGDLHG 161
```

```

Db      | 21 KRVILKALHMK--GRLEESVALRIITEGASILTEQKRLDIDA-----PVTVCSDING 74
|      | 162 KLDLLVLAHKNGLPSSSNPYVNGDFVDRGRKGLLEVLLLSLYLAFNPAFLRNGNE 221
|      | 75 QPFDLMKLEFVGGSSPANTR-YLFLGDYVDRGYFSICVLYMALKILYKTLFLLRNGNE 133
|      | 222 DSVNARVGFIREVESKYPNRNKRILAFIDEVYRWLPGLSVLSRVLVHGGS-DSTSL 280
|      | 134 CSHLTFEYFFKQECIKY---SERVYDACMDAFDCPLAALNQOFLCVHGGLSPEINTL 190
|      | 281 DLKSIDRGKYSILRPPPLTDSBPDLKTEWQOIFDIMSDDPOATWG-----CVPNTLR 333
|      | 191 DDIRKIDRPF-----EPP-----AYGPMCDILMSDPLEDFGNKTOEHFTNTVR 235
|      | 334 GAGWFGPVDVTNPLORHRLSVYIRSHCKPNGHEFMHDK-----ITTPSASNYVAI 387
|      | 236 GGSYFYSYPAVCEFLQHNHNLSTLRAHEAODAGYRMYRSQTTGPPSLITTPSAPYTLV 295
|      | 388 GSNKGAYIRLNNQLM-----PH 404
|      | 296 YNNKAVALKYENNVMMIROPNCSPH 320

```

```

RESULT 23
US-09-744-016A-6
/ Sequence 6, Application US/09744016A
/ Patent No. 6875581
/ GENERAL INFORMATION:
/ APPLICANT: Dr. Voelkel, Helge
/ TITLE OF INVENTION: Method for screening of modulators of calcineurin
/ FILE REFERENCE: A34157PCT
/ CURRENT APPLICATION NUMBER: US/09/744, 016A
/ PRIOR FILING DATE: 2001-10-03
/ PRIOR APPLICATION NUMBER: EP98113876
/ NUMBER OF SEQ ID NOS: 35
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 6
/ LENGTH: 510
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-744-016A-6

```

```

Query Match      11.4%; Score 391.5; DB 2; Length 510;
Best Local Similarity 31.1%; Pred. No. 6.4e-29;
Matches 101; Conservative 58; Mismatches 119; Indels 47; Gaps 10;

|      | 102 KKHIDLLIDVFRKGRGNRLHPKYVALILREAAKSLKQLNPISPVSTAVSQOYTVCGDHLG 161
|      | 45 KRVILKALHMK--GRLEESVALRIITEGASILTEQKRLDIDA-----PVTVCSDING 98
|      | 162 KLDLLVLAHKNGLPSSSNPYVNGDFVDRGRKGLLEVLLLSLYLAFNPAFLRNGNE 221
|      | 99 QPFDLMKLEFVGGSSPANTR-YLFLGDYVDRGYFSICVLYMALKILYKTLFLLRNGNE 157
|      | 222 DSVNARVGFIREVESKYPNRNKRILAFIDEVYRWLPGLSVLSRVLVHGGS-DSTSL 280
|      | 158 CSHLTFEYFFKQECIKY---SERVYDACMDAFDCPLAALNQOFLCVHGGLSPEINTL 214
|      | 281 DLKSIDRGKYSILRPPPLTDSBPDLKTEWQOIFDIMSDDPOATWG-----CVPNTLR 333
|      | 215 DDIRKIDRPF-----EPP-----AYGPMCDILMSDPLEDFGNKTOEHFTNTVR 259
|      | 334 GAGWFGPVDVTNPLORHRLSVYIRSHCKPNGHEFMHDK-----ITTPSASNYVAI 387
|      | 260 GGSYFYSYPAVCEFLQHNHNLSTLRAHEAODAGYRMYRSQTTGPPSLITTPSAPYTLV 319
|      | 388 GSNKGAYIRLNNQLM-----PH 404
|      | 320 YNNKAVALKYENNVMMIROPNCSPH 344

```

```

RESULT 24
US-09-538-092-1283
/ Sequence 1283, Application US/09538092
/ Patent No. 6753314
/ GENERAL INFORMATION:
/ APPLICANT: Manafield, Traci A.
/ TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
/ FILE REFERENCE: 15966-542
/ CURRENT APPLICATION NUMBER: US/09/538, 092
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: 60/127,352
/ PRIOR FILING DATE: 1999-04-01
/ PRIOR APPLICATION NUMBER: 60/178,965
/ PRIOR FILING DATE: 2000-02-01
/ NUMBER OF SEQ ID NOS: 1387
/ SOFTWARE: CuratSeqFormatter Version 0.9
/ SEQ ID NO 1283
/ LENGTH: 521
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (0)...(0)
/ OTHER INFORMATION: Polypeptide Accession Number Q08209
US-09-538-092-1283

```

```

Query Match      11.4%; Score 391.5; DB 2; Length 521;
Best Local Similarity 31.1%; Pred. No. 6.6e-29;
Matches 101; Conservative 58; Mismatches 119; Indels 47; Gaps 10;

```

```

|      | 102 KKHIDLLIDVFRKGRGNRLHPKYVALILREAAKSLKQLNPISPVSTAVSQOYTVCGDHLG 161
|      | 40 KRVILKALHMK--GRLEESVALRIITEGASILTEQKRLDIDA-----PVTVCSDING 93
|      | 162 KLDLLVLAHKNGLPSSSNPYVNGDFVDRGRKGLLEVLLLSLYLAFNPAFLRNGNE 221
|      | 94 QPFDLMKLEFVGGSSPANTR-YLFLGDYVDRGYFSICVLYMALKILYKTLFLLRNGNE 152
|      | 222 DSVNARVGFIREVESKYPNRNKRILAFIDEVYRWLPGLSVLSRVLVHGGS-DSTSL 280
|      | 153 CSHLTFEYFFKQECIKY---SERVYDACMDAFDCPLAALNQOFLCVHGGLSPEINTL 209
|      | 281 DLKSIDRGKYSILRPPPLTDSBPDLKTEWQOIFDIMSDDPOATWG-----CVPNTLR 333
|      | 210 DDIRKIDRPF-----EPP-----AYGPMCDILMSDPLEDFGNKTOEHFTNTVR 254
|      | 334 GAGWFGPVDVTNPLORHRLSVYIRSHCKPNGHEFMHDK-----ITTPSASNYVAI 387
|      | 255 GGSYFYSYPAVCEFLQHNHNLSTLRAHEAODAGYRMYRSQTTGPPSLITTPSAPYTLV 314
|      | 388 GSNKGAYIRLNNQLM-----PH 404
|      | 315 YNNKAVALKYENNVMMIROPNCSPH 339

```

```

RESULT 25
US-09-744-016A-21
/ Sequence 21, Application US/09744016A
/ Patent No. 6875581
/ GENERAL INFORMATION:
/ APPLICANT: Dr. Voelkel, Helge
/ TITLE OF INVENTION: Method for screening of modulators of calcineurin
/ FILE REFERENCE: A34157PCT
/ CURRENT APPLICATION NUMBER: US/09/744, 016A
/ PRIOR FILING DATE: 2001-10-03
/ PRIOR APPLICATION NUMBER: EP98113876
/ NUMBER OF SEQ ID NOS: 35
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 21

```











Db 44 VLKNNH-----VKEG-RVDEEIALRLINEGAALIR-----EKTMIIEBAPITVCGDI 90  
Qy 160 HGLDLDLVVLHKNGLPSSNPVFNQDFVDRGKGLIEVLLLSLYLAFNNAVFLNRGN 219  
Db 91 HGFPLMLKLFVGGSPANTR-YLFLGDDYDRGYFSIECVLYIMVUKILYPSLFLNRGN 149  
Qy 220 HEDSVNNAARYGPIREVSXKPRNKKIILAFIDEVYMLPLGSVLSRVLIVHGGFS-DST 278  
Db 150 HECRHILTEYTFKQCKIKY---SERVYEAACMEAPSLPLAALLNQOFLCVHGLSPRIH 206  
Qy 279 SLDLKSIDRGKVSILRPLPLDGEPLDKTEWQIIPDIMSDPOATWG-----CVPT 331  
Db 207 TLDDIRRLDRFK-----EPP-----AFGPMCDLMSDPSBDFGNEKSGEHSHT 251  
Qy 332 LRAGAVFPGDVTDFLQHRRLSYVIRSHCKPENGHEFMHDK-----ITTFASANY 385  
Db 252 VRGCSYFYNPVAVCEFLQNNNLSTIRAHQAQAGYRMTRKSGTGFPSLITTFSAPLY 311  
Qy 386 AIGSNKGYIRLNNOLM-----PH 404  
Db 312 DVYNNKAVALKYENNVNIRQFNCSPH 338

## RESULT 36

US-08-503-226B-7  
; Sequence 7, Application US/08503226B  
; Patent No. 5871945  
; GENERAL INFORMATION:  
; APPLICANT: Lockebie, Robert Owen, et al.  
; TITLE OF INVENTION: Calcineurin Inhibitory Compounds and  
; TITLE OF INVENTION: Anchoring Protein  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sears Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/503,226B  
; FILING DATE:  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/404,731  
; FILING DATE: 15-MAR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/344,227  
; FILING DATE: 23-NOV-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: 38,659  
; REFERENCE/DOCKET NUMBER: 27866/32861  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-4740448  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 487 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULAR TYPE: protein  
US-08-503-226B-7

Query Match 11.0%; Score 380; DB 1; Length 487;  
Best Local Similarity 30.9%; Pred. No. 7,7e-28;  
Matches 101; Conservative 57; Mismatches 115; Indels 54; Gaps 11;

Qy 100 IRKNHIDLDIVFRKKRGRLHPKYVALIRBAKSLKQLPNISPVSTAVSQOVTGDI 159  
Db 44 VLKNNH-----VKEG-RVDEEIALRLINEGAALIR-----EKTMIIEBAPITVCGDI 90  
Qy 160 HGLDLDLVVLHKNGLPSSNPVFNQDFVDRGKGLIEVLLLSLYLAFNNAVFLNRGN 219  
Db 91 HGFPLMLKLFVGGSPANTR-YLFLGDDYDRGYFSIECVLYIMVUKILYPSLFLNRGN 149  
Qy 220 HEDSVNNAARYGPIREVSXKPRNKKIILAFIDEVYMLPLGSVLSRVLIVHGGFS-DST 278  
Db 150 HECRHILTEYTFKQCKIKY---SERVYEAACMEAPSLPLAALLNQOFLCVHGLSPRIH 206  
Qy 279 SLDLKSIDRGKVSILRPLPLDGEPLDKTEWQIIPDIMSDPOATWG-----CVPT 331  
Db 207 TLDDIRRLDRFK-----EPP-----AFGPMCDLMSDPSBDFGNEKSGEHSHT 251  
Qy 332 LRAGAVFPGDVTDFLQHRRLSYVIRSHCKPENGHEFMHDK-----ITTFASANY 385  
Db 252 VRGCSYFYNPVAVCEFLQNNNLSTIRAHQAQAGYRMTRKSGTGFPSLITTFSAPLY 311  
Qy 386 AIGSNKGYIRLNNOLM-----PH 404  
Db 312 DVYNNKAVALKYENNVNIRQFNCSPH 338

## RESULT 37

US-08-721-458B-7  
; Sequence 7, Application US/08721458B  
; Patent No. 6107104  
; GENERAL INFORMATION:  
; APPLICANT: Lockebie, Robert Owen, et al.  
; TITLE OF INVENTION: Calcineurin Inhibitory Compounds  
; TITLE OF INVENTION: and Anchoring Protein  
; NUMBER OF SEQUENCES: 74  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sears Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/721,458B  
; FILING DATE:  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/404,731  
; FILING DATE: 15-MAR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/344,227  
; FILING DATE: 23-NOV-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: 38,659  
; REFERENCE/DOCKET NUMBER: 27866/33276  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-4740448  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 487 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULAR TYPE: protein  
US-08-721-458B-7





**THIS PAGE BLANK (USPTO)**



Db 121 VSTAVSQVTCGDHKGKDLLLVHKGGLPSSSNPYFNGDFVDRGKGLFVLLLLLS 180  
Qy 205 LVLAPNAVFLNRGNHDSVMMARVGFIRVESKYPNNKRLLAFIDERYRLPLGASVLA 264  
Db 181 LVLAPNAVFLNRGNHDSVMMARVGFIRVESKYPNNKRLLAFIDERYRLPLGASVLA 240  
Qy 265 SHVLVHGGFSDSTSDLLIKSIDRGKYVILRPPLTDGEPDXTKEMQOIFDIIMSDPOAT 324  
Db 241 SHVLVHGGFSDSTSDLLIKSIDRGKYVILRPPLTDGEPDXTKEMQOIFDIIMSDPOAT 300  
Qy 325 MCVPTLIRGAGVWFGPDVTNDFLQHRRLSVYIRSHCKPNGHEFMDNKKITIFASASY 384  
Db 301 MCVPTLIRGAGVWFGPDVTNDFLQHRRLSVYIRSHCKPNGHEFMDNKKITIFASASY 360  
Qy 385 YVIGSNKGAIVIRLNNLMHPVOYISAASQTKLS-KQRMGIYESGALKELAVRMDDHD 444  
Db 361 YVIGSNKGAIVIRLNNLMHPVOYISAASQTKLS-KQRMGIYESGALKELAVRMDDHD 420  
Qy 445 LEDERFRKYDPKDSGYISISHMCKWMENTYKGLPWRLLRDKLAPGTSQKYNVNTL 504  
Db 421 LEDERFRKYDPKDSGYISISHMCKWMENTYKGLPWRLLRDKLAPGTSQKYNVNTL 480  
Qy 505 LDTDVILBAADQMSVMDALYANKASLVAFNIIIDANSGETLDEFEFAIDLVAHMPG 564  
Db 481 LDTDVILBAADQMSVMDALYANKASLVAFNIIIDANSGETLDEFEFAIDLVAHMPG 540  
Qy 565 AYSKAMLEKCMMDLNGCKVDLNEFLFAFRLSDHREKODENIRRRSTGRPSVAKTA 624  
Db 541 AYSKAMLEKCMMDLNGCKVDLNEFLFAFRLSDHREKODENIRRRSTGRPSVAKTA 600  
Qy 625 TDPVTLLADKISNTLVVHEDIDPTDCSRSKVIDPKKS 661  
Db 601 TDPVTLLADKISNTLVVHEDIDPTDCSRSKVIDPKKS 637

RESULT 2  
US-10-369-493-5651  
Sequence 5651, Application US/10369493  
Publication No. US20030233675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
PRIOR FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 5651  
LENGTH: 722  
TYPE: prt  
ORGANISM: Caenorhabditis elegans  
US-10-369-493-5651

Query Match 37.9%; Score 1305; DB 4; Length 722;  
Best Local Similarity 41.6%; Pred. No. 8 3e-99;  
Matches 266; Conservative 117; Mismatches 201; Indels 56; Gaps 13;  
Qy 4 NAIRAFIFLOKYYRRHQAREMQRNMQIFONTLEYASBQDAE-----LYKFNDLIK 57  
Db 90 STIKSAIILIOKYYRRCEARLBARRATWQIFALAFYAGSDQLKRVWVITLDFADVIR 149  
Qy 58 HHPQAAAGRGYQG--SAIVSYLDD--KDYVBERFDYNAKIR----- 97  
Db 150 AAEENGGKGVNGRNSPLMSALSHYAKSLMDSBGETYKMLBETSPNVDIDBNYKGP 209  
Qy 98 ---LPIRKVHIDLLIDVFRKKGGRNLHPKYVALILBAKSLKQJPNISFVSTAVSQVT 154

Db 210 TSLSELDNRQVAKMIEAF--KYNKVLHPKTYVLLIHEARKIKRAMPVSRISTISNOPT 267  
Qy 155 VCGDHLHGKIDLLLVHAKNGGLPSSSNPYFNGDFVDRGKGLFVLLLLSLYLAPNAVF 214  
Db 268 ICGDHLHGKIDLLIILKNGGYSVNPHYFNGDFVDRGQGSIEVLCVLPALVYVDPMSY 327  
Qy 215 LNRGNHDSVMMARVGFIRVESKYPNNKRLLAFIDERYRLPLGASVLAHGGF 274  
Db 328 LNRGNHDSVMMARVGFIRVESKYPNNKRLLAFIDERYRLPLGASVLAHGGF 387  
Qy 275 SDSTLIDLKSIDRGKYVILRPPLTDGEPDXTKEMQOIFDIIMSDPOATMGC 327  
Db 388 SDSTLIDLKSIDRGKYVILRPPLTDGEPDXTKEMQOIFDIIMSDPOATMGC 447  
Qy 328 VENTLIRGAGVWFGPDVTNDFLQHRRLSVYIRSHCKPNGHEFMDNKKITIFASASY 387  
Db 448 VENTLIRGAGVWFGPDVTNDFLQHRRLSVYIRSHCKPNGHEFMDNKKITIFASASY 507  
Qy 388 GSNKGAIVIR-LNNLMHPVOYISAASQTKLS-FKQRMGIYESGALKELAVRMDDHD 445  
Db 508 GSNKGAIVIRFQKSRQPHVQYM--ASKTRKSTLRERLGVYESAAYELKSKLSFHD 565  
Qy 446 LEDERFRKYDPKDSGYISISHMCKWMENTYKGLPWRLLRDKLAP-GTDSQKYNVNTL 504  
Db 566 LOKERFINDIKSGKPLIKMSDCYERITGLPPIALAPKVAITLSDGKYVNYKEDRI 625  
Qy 505 LDTDVILBAADQMSVMDALYANKASLVAFNIIIDANSGETLDEFEFAIDLVAHMPG 564  
Db 626 AQVGTHAQEKD---IVSLVHNRKSTLTFTLFFMDKONGQVSMKEBFDACVLT-----G 677  
Qy 565 AYSKAMF-----LEKCMMDLNGCKVDLNEFLFAFRLSD 599  
Db 678 KYTKRPLQTDYISQIAESIDFNKDGFDLNELEAFRLVD 717

RESULT 3  
US-10-287-226-324  
Sequence 324, Application US/10287226  
Publication No. US20040086875A1  
GENERAL INFORMATION:  
APPLICANT: Agee, Michele L.,  
APPLICANT: Alsbrook, John P.,  
APPLICANT: Berghs, Constance,  
APPLICANT: Boldog, Ference,  
APPLICANT: Burgess, Catherine E.,  
APPLICANT: Chant, John S.,  
APPLICANT: Chaudhuri, Amitabha,  
APPLICANT: Dipippo, Vincent A.,  
APPLICANT: Edinger, Shlomit R.,  
APPLICANT: Eisen, Andrew,  
APPLICANT: Ellerman, Karen,  
APPLICANT: Gargolli, Sasha A.,  
APPLICANT: Gorman, Linda,  
APPLICANT: Gerlach, Valerie,  
APPLICANT: Uli, Weizhen,  
APPLICANT: Kekuda, Ramesh,  
APPLICANT: Khramtsov, Nikolai,  
APPLICANT: Li, Li,  
APPLICANT: Malvankar, Uriel M.,  
APPLICANT: MacDougall, John R.,  
APPLICANT: Mezes, Peter S.,  
APPLICANT: Miller, Charles E.,  
APPLICANT: Millet, Isabelle,  
APPLICANT: Ooi, Chean Eng,  
APPLICANT: Ort, Tatiana,  
APPLICANT: Padigaru, Muralidhara,  
APPLICANT: Paturajan, Meera,  
APPLICANT: Rastelli, Luca,  
APPLICANT: Rieger, Daniel K.,  
APPLICANT: Rothenberg, Mark E.,  
APPLICANT: Shenoy, Suresh G.,  
APPLICANT: Spaderna, Steven K.,



```

: APPLICANT: Spytek, Kimberley A.,
: APPLICANT: Taupier, Jr., Raymond J.,
: APPLICANT: Vermet, Corine A.M.,
: APPLICANT: Zernhusen, Bryan D.,
: APPLICANT: Zhong, Mei
: TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
: FILE REFERENCES: 21402-480C
: CURRENT APPLICATION NUMBER: US/10/287,226
: CURRENT FILING DATE: 2002-11-04
: PRIOR APPLICATION NUMBER: 60/334,421
: PRIOR FILING DATE: 2001-11-30
: PRIOR APPLICATION NUMBER: 60/354,392
: PRIOR FILING DATE: 2002-02-04
: PRIOR APPLICATION NUMBER: 60/360,148
: PRIOR FILING DATE: 2002-02-27
: PRIOR APPLICATION NUMBER: 60/364,000
: PRIOR FILING DATE: 2002-03-13
: PRIOR APPLICATION NUMBER: 60/404,821
: PRIOR FILING DATE: 2002-08-20
: PRIOR APPLICATION NUMBER: 60/334,526
: PRIOR FILING DATE: 2001-11-30
: PRIOR APPLICATION NUMBER: 60/354,409
: PRIOR FILING DATE: 2002-02-04
: PRIOR APPLICATION NUMBER: 60/364,227
: PRIOR FILING DATE: 2002-03-13
: PRIOR APPLICATION NUMBER: 60/334,027
: PRIOR FILING DATE: 2001-11-28
: PRIOR APPLICATION NUMBER: 60/331,641
: PRIOR FILING DATE: 2001-11-20
: Remaining Prior Application data removed - See file Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 673
: SOFTWARE: Cnaseqblast version 0.1
: SEQ ID NO 324
: LENGTH: 613
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-10-287-226-324

Query Match      33.8%; Score 1164.5; DB 4; Length 613;
Best Local Similarity 39.3%; Pred. No. 2,8e-87;
Matches 247; Conservative 115; Mismatches 200; Indels 67; Gaps 13;

QY      3  ENAIAATFIQKMYRRHQARREORRCMNQIFQNLLEYASEQDAELHYKFPNDLIKMPQA 62
DB      14  DTSLRALITIQWYRGYKARKLAKROYALTITPOSLEYADBOGQMOISTEFSFLEUYTHI 73
QY      63  AGRKNOYGSAHVSVLDDKD--DLVE--EFGDIVNA-KIELPRKXHIIDLLIDVFKKGG 117
DB      74  HKEBELRNQSLSEBQDMQDRWDYDSDIVPDSYNGPRLOPFLCTCIDLLLEAFKQO- 132
QY      118  NLLHPKYVALILREAAKSLIKQLPNISPVSTAVSQOYTVCGDLHGKLDLLVVLHKKGLPS 177
DB      133  -ILHAYVLEVLFETKQVKLQKMPNFTHIQTSPSKEVTICGDLHGKLDLFLFYKNGGLPS 191
QY      178  SNNPVYENDFYDRGKRGLEVLILLSLTLPAPNAVFLRGNHEDSVMMARCFIREVS 237
DB      192  EBNPVYENDFYDRGKNSLEILMLTLCVSLPYVNDLIRGNHEDPMNLRGFTPEILH 251
QY      238  KTPRNHRIILAFIDEVYRMLP-----LGSVLNRSVLIVHGFSDSTSLD 282
DB      252  KYKLGKRILOLLEBYVALPTEITNRDHDGDSKNNKGVTFPA-----HGRIK----- 299
QY      283  IKSIRGKVTSLRPELTGEP--LDKTEWQOIFDIWMSDPOATMGCVPTLIRGAWPF 339
DB      300  -----TNGSPTEHTEHEWQOIIDILMSDRGKNGCFPTCRGGGCVF 342
QY      340  GPDVTNPLQHRHLSVISHCKPGHGFPMDNKTIITFSASNYATISNGAYIRLNN 399
DB      343  GPDVTSKILNKYQLKWLIRSHCKPGEYBICHDGKVVITFSASNYEBSNGAYIKLCS 402
QY      400  QLMPHFVOYISAASQTKRLSF--KORMGIVESSALKELAVMMRDHRELBDPFRKYDPK 456
DB      403  GTTPRFQY-----QYTKATCFQPLRQRYVTMENSALKILREKIVISKSLTPTAFQLODHR 458

```

```

QY      457  DSGYISISWCKMYMENTGLGPMRLLRDKLAPGTDSDQVNRRTDLDLTD-VIIRAEA 515
DB      459  KSGKLSVSQWACMENNILGILNFWRLSSNLVNIIDONGVYVWSSPQNRIRKPVQBAHS 518
QY      516  DGMSVNDALYANKASIVAFENIIDANSGETLTDEPETAIDLLVAMHPGAYSAAEMLEKC 575
DB      519  ---TLVEITRYRSDLBITFNALIDTHSGLSIVERRAMMKLFSSHYNVHIDDSQVKNKLA 575
QY      576  RMDLNGDGKVDLNEFLAEFRUSDLRKE 604
DB      576  NIMDLNKGSDIDPNEFLKAFYV--VARYE 602

RESULT 4
US-10-287-226-326
: Sequence 326, Application US/10287226
: Publication No. US20040086875A1
: GENERAL INFORMATION:
: APPLICANT: Agee, Michele L.,
: APPLICANT: Alsbrook, John P.,
: APPLICANT: Berghs, Constance,
: APPLICANT: Boldog, Ference,
: APPLICANT: Burgess, Catherine E.,
: APPLICANT: Chant, John S.,
: APPLICANT: Chaudhuri, Amitabha,
: APPLICANT: DiIppio, Vincent A.,
: APPLICANT: Edinger, Shlomit R.,
: APPLICANT: Elsen, Andrew,
: APPLICANT: Ellerman, Karen,
: APPLICANT: Gangolli, Esha A.,
: APPLICANT: German, Linda,
: APPLICANT: Gerlach, Valerie,
: APPLICANT: Ji, Weizhen,
: APPLICANT: Kekuda, Ramesh,
: APPLICANT: Khamteev, Nikolai,
: APPLICANT: Li, Li,
: APPLICANT: Malyankar, Uriel M.,
: APPLICANT: MacDougall, John R.,
: APPLICANT: Mezes, Peter S.,
: APPLICANT: Miller, Charles E.,
: APPLICANT: Millet, Isabelle,
: APPLICANT: Ooi, Chean Eng,
: APPLICANT: Ort, Tatiana,
: APPLICANT: Padigaru, Muralidhara,
: APPLICANT: Patcureajan, Meera,
: APPLICANT: Rieger, Daniel K.,
: APPLICANT: Rothenberg, Mark E.,
: APPLICANT: Shenoy, Suresh G.,
: APPLICANT: Spaderma, Steven K.,
: APPLICANT: Spytek, Kimberley A.,
: APPLICANT: Taupier, Jr., Raymond J.,
: APPLICANT: Vermet, Corine A.M.,
: APPLICANT: Zernhusen, Bryan D.,
: APPLICANT: Zhong, Mei
: TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
: FILE REFERENCES: 21402-480C
: CURRENT APPLICATION NUMBER: US/10/287,226
: CURRENT FILING DATE: 2002-11-04
: PRIOR APPLICATION NUMBER: 60/334,421
: PRIOR FILING DATE: 2001-11-30
: PRIOR APPLICATION NUMBER: 60/354,392
: PRIOR FILING DATE: 2002-02-04
: PRIOR APPLICATION NUMBER: 60/360,148
: PRIOR FILING DATE: 2002-02-27
: PRIOR APPLICATION NUMBER: 60/364,000
: PRIOR FILING DATE: 2002-03-13
: PRIOR APPLICATION NUMBER: 60/404,821
: PRIOR FILING DATE: 2002-08-20
: PRIOR APPLICATION NUMBER: 60/334,526
: PRIOR FILING DATE: 2001-11-30
: PRIOR APPLICATION NUMBER: 60/354,409

```



```
QY 105 IDLLIDVFRKRGNRRLHPKRYVALILREAAKSLKOLPNISPVSTAVSQOVTVCGLHGKLDL 164
Db 181 VKAMLEBFKQK--CIHKRYAYQIVLTKTLLLSMSLSLVDDVDPNGCHFTVCGVMQYF 238
QY 165 DLVLVHLKNGLPSSSNPVYFNGDFVDRGKRGLEVLILLLSLYLAFPAVAFLNRGNHDSY 224
Db 239 DLINIPDNLGSPSEBNRYLFGNGDFVDRGSPSVAYILTLFAFKCLYPRAMTLAAGNHSKS 238
QY 225 MNAHYGRIREVESKYPNNHRIALFIDVRYRWLPGLSVNLSRYLVHGGF--PS-DSTSLDL 282
Db 299 MNKTYGEGEVRKDLGKFPVLPFA---EVPCLPLAVINNKVYVHGGFSPYDGVGLSD 355
QY 283 IKSIDRGKYVSLRPPLTGDGRLDKTEWQOIFDIIMSDDPQATWGCVNPLTRGAWMPGPD 342
Db 356 IRSIDR-----FCBPP-----EGLMCEILMSDPQOLGKPSK--RGVALSFGAD 399
QY 343 VTDFLQRLSLVYIRSHCKPNHGFMDNKLITTFASNYAIGSNKAYIRLN--NQL 401
Db 400 VTKKFLQENNLDLIVRSHVYKDSGHELEHDKLITVSAPNYCDQMNKGAFTIRTAPEM 459
QY 402 MHPFYQY 408
Db 460 KPDIYTF 466

RESULT 7
US-11-097-143-18312
; Sequence 18312, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18312
; LENGTH: 520
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-18312

Query Match 16.8%; Score 577; DB 6; Length 520;
Best Local Similarity 32.1%; Pred. No. 1.1e-38;
Matches 136; Conservative 83; Mismatches 153; Indels 52; Gaps 15;

QY 4 NAIRA-AIFQKTYRRHQAREM-----QRGCMQIFQNLRYASQDAQLYKFFNDLIK- 57
Db 108 SAYADADYALKGYRRAAHMSLGKFRQALCDPEFAVKCR--PNDKAKLFTBCNKLVKM 166
QY 58 -----HMPQAGRKNOYQGSAAHVSLLD--KDDVSEFGDIVNAKIELPIRKNIIDL 107
Db 167 RAPEBALAVKPEKT-LSBNYSMDMENTITIEDYKGPQLED-----GKVLTKPKKE----- 215
```

```
QY 108 LIDVFRKRGNRRLHPKRYVALILREAAKSLKOLPNISPVSTAVSQOVTVCGLHGKLDL 167
Db 216 LMEHYRAQK--RLHRRFAVKILICEIDTYMRAQPSLVDITVPEBEKFTICGDHGOFPYDLM 273
QY 168 VVLVHLKNGLPSSSNPVYFNGDFVDRGKRGLEVLILLLSLYLAFPAVAFLNRGNHDSYMA 227
Db 274 NTFEINGLPSKPNFYLFNGDFVDRGSPSVAYILTLFAFKCLYPRAMTLAAGNHSKINMO 333
QY 228 RYGFIREVSKYPNNHRIALFIDVRYRWLPGLSVNLSRYLVHGGF--SDSTSLDLKS 285
Db 334 MYGFTGEVTAQKT---SAMADIFQYFNNLPLCHCINQKILVHGGFSTEDVTLDHIR 390
QY 286 IDRGKYVSLRPPLTGDGRLDKTEWQOIFDIIMSDDPQATWGCVNPLTRGAWMPGPDYTD 345
Db 391 IERN-----CQPP-----EGLMCEILMSDPQOMWG--LGQSRGVGIQFGPDYTR 434
QY 346 NPLQRLSLVYIRSHCKPNHGFMDNKLITTFASNYAIGSNKAYIRLN--NQLMPH 404
Db 435 KFGKNNLDYIIRSHVYKDSGHELEHDKLITVSAPNYCDTGMNKAFTITGNLKEPN 494
QY 405 FVQY 408
Db 495 YKSP 498
```

```
RESULT 8
US-10-450-763-40348
; Sequence 40348, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 40348
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (405)..(460)
; OTHER INFORMATION: Serine/threonine specific protein phosphatases proteins
; OTHER INFORMATION: domain identified by eMATRIX, accession number BL00125D, p-value=
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (207)..(468)
; OTHER INFORMATION: Ser/Thr protein phosphatase domain identified by Pfam,
; OTHER INFORMATION: accession name STHOSPHATASE, E-value=1.4e-106, Pfam score of 3.67
; NAME/KEY: misc_feature
; LOCATION: (1)...(500)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-40348

Query Match 16.6%; Score 571; DB 5; Length 500;
Best Local Similarity 34.8%; Pred. No. 3.1e-38;
Matches 143; Conservative 73; Mismatches 153; Indels 42; Gaps 16;
```

```
QY 11 FIOKTYRRHQAREMQR--RCNMQIFQNLRYAS--BDDQALYKFFNDLIKHMPOAAGRN 67
Db 99 YIMGYRRAASNNALGKFPALRLDYEYTVKVPKPHDADAKYQDECNKIYK--PKAPER-- 154
```

QY 68 OYOGSAHV-SVLDKX-----DLVEBFG--DIVAKIELPIRKHIDLIDVERKKRGNL 120  
DB 155 ALAGDBHKRSVDSLDISEMTIEDYSXKLEBGKTIIPMK-----LMQWKXK--KL 208  
QY 121 HPTVVALIIREAAKSLKQLEPNISPVSTAVSQOVTGCDLHGKLDLLVLAHNGLPSSN 180  
DB 209 HRCKAQOILVQVEKVLSTKSTLVETTLKETEKITVCGDTHGQFYDLNITFELNGLPSEFN 268  
QY 181 PVYENGDFVDRGRGLGVLLLLSLYAFPNVFLNRGNHDSVMNARGFIREVSKRP 240  
DB 269 PFIENDFVDRGSPSVSVLLTLPFGFLLYPDHFHLNRGNHEDVMNQIFGEGEVAKYT 328  
QY 241 RHRKRIAFIDEVYRWLPFGSVLNSRVLIVHGPF--SDSTSLDLIKSIDRGKVSILRPP 298  
DB 329 ---AGMYELFSVFEVFLPLAQCGINGVLLMHGGLFSEBDVTLDDIRKIERNR-----QP 380  
QY 299 LIDGEPDLKTEWQOIFDIWMSDPOATMGCVPTLRGAGVFGPDVTNFIQRHRLSYIR 358  
DB 381 --DSGPM-----CDLMSDPQPONG--RSISKRVTCQFGPDVTKAFLENNLDYIR 429  
QY 359 SHECKNGHEFMHNDKIIITIFASANYAIGSNKGAIVRL--NQMLMHPVOY 408  
DB 430 SHEVKAEGTEVAHGRCVTVFSAPNYCDQMGNASTIHLQSGDLRPFQHP 480

RESULT 9  
US-10-437-963-195520  
Sequence 195520, Application US/10437963  
Publication No. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
FILE REFERENCE: 38-21(53221)B  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 195520  
LENGTH: 483  
TYPE: PRT  
ORGANISM: *Oryza sativa*  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_91460C.1.pep  
US-10-437-963-195520

Query Match 16.6%; Score 570.5; DB 4; Length 483;  
Best Local Similarity 34.8%; Pred. No. 3.3e-38;  
Matches 145; Conservative 69; Mismatches 154; Indels 49; Gaps 12;

QY 9 AIFIOKMYRRHQRREMR--RCNMQIFOMLEYASBDQAEIYKFNDLIKMPQAAGRKN 67  
DB 77 ARYSKYTRKGAAYLAMGFKKALKDFOQVKRISPNDP-----DARRKICEKAKAV 128  
QY 68 OYOGSAHVSLDKDLDVEBFGDIVAKIEL-----PIRKHIDLIDVERKKRGN 117  
DB 129 KIRFEBALISVGBEKRKSVASIDYRIIEVBPOYTGPRVNDGDTITLDPVAKMLDEFERK 187  
QY 118 NRIHPYVALIIREAAKSLKQLEPNISPVSTAVSQOVTGCDLHGKLDLLVLAHNGLP 177  
DB 188 -CHHKYAVQIVLQTLRSVPSLVDPVDSHFTVCGDVAGQYFDLLNIFKLNGLRS 246  
QY 178 SSGPYFNGDFVDRGRGLGVLLLLSLYAFPNVFLNRGNHDSVMNARGFIREVSKRP 237  
DB 247 EENPYFNGDFVDRGSPSVSVLLTLPFAFKCLVPTGMYLARGNHESNMKIYGFEBEVR 306  
QY 238 KYPRNKRILAFID--EYVRWLPGLSVLNSRVLIVHG--FS-DSTSLDLIKSIDRGKTV 292

DB 307 KUGS-----AFIBAFARFCLPLAHVINNVFVHGGLFSDVDKSLDIRAIR----- 356  
QY 293 SLIRPPLDGEPLDKTEWQOIFDIWMSDPOATMGCVPTLRGAGVFGPDVTNFIQRH 352  
DB 357 --FREPPESG-----LMCEVLSBDPOPLGRGPRK--RGVGLSFGADVTYKCKLQENN 404  
QY 353 LSVYRSHCKNGHEFMHNDKIIITIFASANYAIGSNKGAIVRL--NQMLMHPVOY 408  
DB 405 LDIIVRSHVDRGGEIIEHDKLITVFSAPNYCDQMGNAATIRTPALKEPNISSF 461

RESULT 10  
US-10-732-923-9610  
Sequence 9610, Application US/10732923  
Publication No. US20050108791A1  
GENERAL INFORMATION:  
APPLICANT: Edgerton, Michael D  
TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
FILE REFERENCE: 38-15(52796)C  
CURRENT FILING DATE: 2003-12-10  
PRIOR FILING DATE: 2002-12-04  
NUMBER OF SEQ ID NOS: 24149  
SEQ ID NO 9610  
LENGTH: 485  
TYPE: PRT  
ORGANISM: *Lycopersicon esculentum*  
US-10-732-923-9610

Query Match 16.6%; Score 570.5; DB 5; Length 485;  
Best Local Similarity 36.1%; Pred. No. 3.3e-38;  
Matches 140; Conservative 55; Mismatches 124; Indels 69; Gaps 12;

QY 24 EMQRRCNQIFOMLEYASBDQAEIYKFNDLIKMPQAAGRKNQYOGSAHVSVDKXD 83  
DB 142 ESQRR--SVADSIDYRS-----VEYBPQYAGARIE----- 169  
QY 84 LVEBFGDIVAKIELPIRKHIDLIDVFRKKGRNRLHPKYVALIIREAAKSLKQLEPNIS 143  
DB 170 -----GDVYTL-----DFVKGMLDDEPKNOK--NLHRYAVQIVLQTRRMLRALPSLV 214  
QY 144 PVSTAVSQOVTGCDLHGKLDLLVLAHNGLPSSNPFVFGDPVDRGRGLGVLLLL 203  
DB 215 DIVPBGKHFVCGVHGQFIDLNIIFELNGLPSEDNPLFNGDVPDRGSFSLVLLTF 274  
QY 204 SLTYAFPNVFLNRGNHDSVMNARGFIREVSKYPRNHRKILAFIDEVYRWLPFGSVL 263  
DB 275 AFKMCPSALHILARNHBSKSNKIYGFEGEVRSLSEIFVELFA---EVFCCLPLAHVI 331  
QY 264 NSRVILVHGQ--PS-DSTSLDLIKSIDRGKTVSILRPPDLTGEPDLKTEWQOIFDIWMSDP 321  
DB 332 NKRVVAVHGGLPSSVGVKLSDIRAIR-----FCBPP-----EBGLMCELSMDP 376  
QY 322 GATMGCVPTLRGAGVFGPDVTNFIQRHRLSYIRSHCKNGHEFMHNDKIIITIFSA 381  
DB 377 QPQPRGPRK--RGVGLSFGADVTYKCKLQENNLDIIVRSHVDRGGEIIEHDKLITVESA 435  
QY 382 SNYTAIGSNKGAIVRL--NQMLMHPVOY 408  
DB 436 PNYCDQMGNAATIRTPALKEPNISSF 461

RESULT 11  
US-10-043-487-296  
Sequence 296, Application US/10043487  
Publication No. US20030055220A1  
GENERAL INFORMATION:  
APPLICANT: HYBRIGENICS  
APPLICANT: PIERRE, LEBERAIN  
TITLE OF INVENTION: Protein-protein interactions between Shigella flexneri polypeptide  
TITLE OF INVENTION: mammalian polypeptides



```
Db 374 QPQPGPGRK-RGVGLSGFDVTKRFLQNNLIDLVRSHVXDQGYEIHDKLITVPSA 432
QY 382 SNYYAIGSNKGAAYIRLN-NQLMPHFVQY 408
Db 433 PNYCDQMGKGAFFIRFADPMKNIVTF 460

RESULT 14
US-10-298-638-10
; Sequence 10, Application US/10298638
; Publication No. US20030177527A1
; GENERAL INFORMATION:
; APPLICANT: HARRING, Michael A. et al.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING SIGNAL TRANSDUCTION
; TITLE OF INVENTION: COMPONENTS IN DURABLE AND BROAD-RANGE RESISTANCE
; FILE REFERENCE: 2121-0167P
; CURRENT APPLICATION NUMBER: US/10/298,638
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 10/298,638
; PRIOR FILING DATE: 2002-11-19
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-10-298-638-10

Query Match 16.3%; Score 562.5; DB 4; Length 485;
Best Local Similarity 35.8%; Pred. No. 1.5e-37;
Matches 139; Conservative 55; Mismatches 125; Indels 69; Gaps 12;

QY 24 EQQRRCNQIFQNLSEYASRQDAELKFFNDLIKHMPOAGRKQNGYQSAHVSVDKDD 83
Db 142 EQQR--SVADSIDRS-----VETPQYAGARIR----- 169
QY 84 LVEBFGDIYNAKIELPIRKNIIDLIDVRKKGKGNLHPKYVALILREAAKSLKOLENIS 143
Db 170 -----GDVVTL-----DVGKMLDDPKNQK--NLHGRVAYQIVLQIRMLRALPSLV 214
QY 144 PVTSTASQGVTVCGDLHGKLDLVLVLRKNGLPSSNPYVNGDFYDRGKGLVLLIL 203
Db 215 DIVVEGKGFVCGDVHGQFYDLNLFRLNGLPSEBNDPYLFNGDFYDRGSPSLVLLTLF 274
QY 204 SLYLAFNNAVFLNRGNHDSVNNARVGFIRVEYSKTPRNHKKILAFIDSVYMLPLGSYL 263
Db 275 AKKCMCPSAIHLARGNHESKSMNKIYGFEGEVRSKLSIFVSLFA---EVPCCLPLAHVI 331
QY 264 NSRVLIVHGQ-FS-DSTSLDLIKSIDRGKTVSILRPPLTDGSRPLDKTEHQIFDIWMSDP 321
Db 332 NKKVPLVHRGLFSVQGVKLSIDTRALDR-----FCERP-----EGLMCELLMSDP 376
QY 322 QATMGCVPTLRGAGVGFDPVDTNLFQHRRLSYVIRSHCKPENGHEFMHDKIITITPSA 381
Db 377 QPQPGPGRK-RGVGLSGFDVTKRFLQNNLIDLVRSHVXDQGYEIHDKLITVPSA 435
QY 382 SNYYAIGSNKGAAYIRLN-NQLMPHFVQY 408
Db 436 PNYCDQMGKGAFFIRFADPMKNIVTF 463

RESULT 15
US-10-424-599-191442
; Sequence 191442, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
```

```
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 191442
; LENGTH: 482
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(482)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: PAT_MRT3847_14893C.1.pep
US-10-424-599-191442

Query Match 16.3%; Score 561; DB 4; Length 482;
Best Local Similarity 38.4%; Pred. No. 2e-37;
Matches 135; Conservative 46; Mismatches 119; Indels 52; Gaps 10;

QY 60 PQAARKNGYQSAHVSVDKDDLVESRGDIVNAKIELPIRKNIIDLIDVFRKKGNR 119
Db 158 PQYSGAR--IRGDA--VTLDFVKKMLIDF-----KNQK 187
QY 120 LHPKTVALLREAAKSLKOLPNISPVSTAVSQGVTVCGDLHGKLDLVLVLRKNGLPSSS 179
Db 168 LHKRYAPQIVQTRKXLDALPSLVINHVPDGHFTVCGDVHGQYVDLNIFFELNDLPSE 247
QY 180 NPVYNGDFVDRGKGLVLLILSLYLAFNNAVFLNRGNHDSVNNARVGFIRVEYSKY 239
Db 248 NPVYNGDFVDRGSPSLVLLILFLAFKCMCPSAIHLARGNHESKSMNKIYGFEGEVRSKL 307
QY 240 PRNHKRIIAFDVRYRMLPLGSVLSRVLVHGQ-FS-DSTSLDLIKSIDRGKTVSILRP 297
Db 308 NETFVELFA---EVPCCPLAHVINRKYFVHVGGLFSVDGVKVSIDRSINR-----FCER 359
QY 298 PLTDEPDLKTEWQIFDIWMSDPQATMGCVPTLRGAGVGFDPVDTNLFQHRRLSYVI 357
Db 360 P-----EGLMCELLMSDPQLPGRGPRK-RGVGLSGAGAVTRFLQENNLIDLIV 408
QY 368 RSHECKPENGHEFMHDKIITITFSASNYTYAIGSNKGAAYIRLN-NQLMPHFVQY 408
Db 409 RSHVEYKDGYSRIHDKLITVFSAPNYCDQMGKGAFFIRFADPMKNIVTF 460

RESULT 16
US-10-298-638-28
; Sequence 28, Application US/10298638
; Publication No. US20030177527A1
; GENERAL INFORMATION:
; APPLICANT: HARRING, Michael A. et al.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING SIGNAL TRANSDUCTION
; TITLE OF INVENTION: COMPONENTS IN DURABLE AND BROAD-RANGE RESISTANCE
; FILE REFERENCE: 2121-0167P
; CURRENT APPLICATION NUMBER: US/10/298,638
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 10/298,638
; PRIOR FILING DATE: 2002-11-19
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 533
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-298-638-28

Query Match 16.2%; Score 557.5; DB 4; Length 533;
Best Local Similarity 39.4%; Pred. No. 4.6e-37;
Matches 127; Conservative 55; Mismatches 109; Indels 31; Gaps 10;

QY 94 AKIE-LPIRKNIIDLIDVFRKKGNRLHPKYVALILREAAKSLKOLPNISPVSTAVSQ 152
```

Db 214 ARIEGEVTLDPVKTMEDFNOK--TLHRYAYQVLTQRTLLALPSLVDSVPHGK 271  
Qy 153 VVWCGDLHGKLDLVLVHLHNGKLPSSSNPVYFNGDPDRKRGLEVTLLLSLYLAPNA 212  
Db 272 TVCGDVHGQFYLLMLFBLNGLPSEENPYLFNGDFDRSGFVEILLTLFARKCMPS 331  
Qy 213 VFLRGNHEDSVNARYGFIREVESKYPNNHKLIAFID--EYRMLPLGSVYLSRVLI 269  
Db 332 IYLARKNHESKNNKIYGFEGEVRSLKSEK-----FVLLFAVFCYCLPLAHYINGKVF 385  
Qy 270 VHGG-FS-DSTSLDLKSIDRGKYSTLRPPLTDEBPLDKTEWQOIFDIMSDPOATMG 327  
Db 386 VHGGFLSVGVGKSLDIRAIDR-----FCEBP-----EGLMCELLMSDPQPLPGR 430  
Qy 328 VEPYTLGAGWYFEPDVTNDNLFQHRSLSYVRSHECKNGHEFPHNDKITIFSAASYAI 387  
Db 431 GPEK-RGVGLSFQGDVYTKRFLQNNMLDLVRSHEVNDGVEVHDGDLTVFSAFVYCD 489  
Qy 388 GSNKGAIVIRLN-NQLMPHFVQY 408  
Db 490 MGKGAIFIRFAPDMKFNIVTF 511  
RESULT 17  
US-10-369-493-22595  
; Sequence 22595, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052) B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 22595  
; LENGTH: 473  
; TYPE: PRT  
; ORGANISM: Schizosaccharomyces pombe  
US-10-369-493-22595  
Query Match 15.1%; Score 519; DB 4; Length 473;  
Best Local Similarity 35.2%; Pred. No. 5.9e-34;  
Matches 119; Conservative 66; Mismatches 127; Indels 26; Gaps 8;  
Qy 73 AHVSVLDDKDDLVBEFGDIVNAKIELPIRKNIHLLIDVFRKKRGNRILPKYVALILREA 132  
Db 139 ANINIRD--MDIPSDYDYL--LBKQITKBFVEDMKERF--CQGGKPLPKAYSLIRLG 191  
Qy 133 ASLKLPLNISPSTIVSQOVTYCGDLGKLDLVLVHLHNGKLPSSSNPVYFNGDFVDRG 192  
Db 192 KLELEKTPSLIDIPVKGDELVLVCGPTHGQYFDLNIPLKHGPPSPNTKYLFGGDFVDK 251  
Qy 193 KRGLEVTLLLSLYLAPNAVFLRGNHEDSVNARYGFIREVESKYPNNHKLIAFID 252  
Db 252 SMSTEVAFYLYAKLYPDVFLRGNHETDDNNKYVFGEGCRSKY--NERTFNIPSE 308  
Qy 253 YVRMLPLGSVLSRVLIHGGF--SDSTSLDLKSIDRGKYSTLRPPLTDEBPLDKTEW 310  
Db 309 TSLPLPLGSLISDYSVLVHGGFLPSDDNVTLLDRLNIDRS--KQPGQSG----- 356  
Qy 311 QOIFDIMSDPOATMGCPYNTLRGAGWYFEPDVTNDNLFQHRSLSYVRSHECKNGHEF 370  
Db 357 -LMMELMTDPOAPRGPSK--RGVGLQGPVDSKPFCEANGKAKAVIRSHVADQGYEVE 414  
Qy 371 HDNKITIFSAASYAIGSNKGAIVIRLNQMLMPHFVQY 408

Db 415 HDGCVTFPSAPNYCDSYGNLGAIVIKYKEDMELDFHQF 452  
RESULT 18  
US-10-369-493-4091  
; Sequence 4091, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052) B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 4091  
; LENGTH: 598  
; TYPE: PRT  
; ORGANISM: Neurospora crassa  
; FEATURES:  
; NAME/KEY: unsure  
; LOCATION: (1)..(598)  
; OTHER INFORMATION: unsure at all Xaa locations  
US-10-369-493-4091  
Query Match 14.6%; Score 502.5; DB 4; Length 598;  
Best Local Similarity 29.2%; Pred. No. 2e-32;  
Matches 141; Conservative 55; Mismatches 148; Indels 139; Gaps 12;  
Qy 12 IQKWRBHQARREMRQCMQIFQNLVYASBDDQAEVYFNDLIGMPQAGRKQYOG 71  
Db 158 VAYIRRATA-----YAILNPKAAVDFPTCVIAP----- 189  
Qy 72 SAHVSVLDDKD--DLVE-----EFGDIVNAKIELPIRKNIHLLID-- 110  
Db 190 -----DKDKAKLKVBECEKIVROLAPFAALEVDELSAAGLDVESNAVDSYGVVL 242  
Qy 111 -----VPRKKRGNRLHPKYVALILREAAKSLIKOLPNISPSTAVSQOVTYCG 157  
Db 243 EGNEMTQEFIDMIRFRKRGKLIHKKYVQIILAVANIYVNEPTWVEVDIPEVQLTVCG 302  
Qy 158 DLHG-----KLDLVLVHLHNGKLPSSSNPVYFNGDPDRKRGKLE 197  
Db 303 DTHGMPALAMPSPSKWLLTNCEKAYQFDLMELFRLNGFPSPDKRYLLFNGDFVDRGWS 362  
Qy 198 VLLLSLYLAPNAVFLRGNHEDSVNARYGFIREVESKY----- 239  
Db 363 IALLLYAYKMLRPNRPFIRGNHETDDNNKYVFGEGCHAKNNEKASTLCAHYATSKI 422  
Qy 240 -----PRNHKRIIAFIDEVYRMLPLGSVLSRVLIHGGF--SDSTSLDLK 285  
Db 423 PHTPADTSHLLPRTYK-----LFSSEFSLPLATILIGKFLVHGGFLPSDDNVTLLD 478  
Qy 286 IIRGKYVSLRPPPLTDEBPLDKTEWQOIFDIMSDPOATMGCPYNTLRGAGWYFEG 345  
Db 479 LDRHKO--KQPGQSG-----LMMELMTDPOAPRGPSK--RGVGMQFGPDVTK 524  
Qy 346 NFLQHRSLSYVRSHECKNGHEFPHNDKITIFSAASYAIGSNKGAIVIRLNQMLMPHF 405  
Db 525 RFCDNKGLAIIIRSHVADQGYEVEHHDGKCTVFSAPKCDMTEKGAIIINIGPYKLF 584  
Qy 406 VOY 408  
Db 585 SQF 587  
RESULT 19

US-10-450-763-40347  
; Sequence 40347, Application US/10450763  
; Publication No. US20050196754A1  
; GENERAL INFORMATION:  
; APPLICANT: HySeq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/549,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 40347  
; LENGTH: 714  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (356)..(384)  
; OTHER INFORMATION: SERINE/THREONINE PHOSPHATASE FAMILY SIGNATURE domain  
; OTHER INFORMATION: identified by eMATRIX, accession number PR00114A, p-value=4.600e-  
; OTHER INFORMATION: 25, raw score of 14.44  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (321)..(696)  
; OTHER INFORMATION: Ser/Thr protein phosphatase domain identified by Pfam,  
; OTHER INFORMATION: accession name Stposphatase, E-value=4.6e-92, Pfam score of 319,  
US-10-450-763-40347

Query Match 14.4%; Score 496.5; DB 5; Length 714;  
Best Local Similarity 29.2%; Pred. No. 8e-32;  
Matches 140; Conservative 62; Mismatches 147; Indels 131; Gaps 15;  
41 SSODAEIKYFFNDLIKHPQAGRNQYQ-----GSAHV-SYLDXQ- 82  
231 SEHEATLTLV---VVKVPHDKDAKMKQSCNKIVKQAFBRALAGDEHKRSVDSLDI 287  
83 ---DLVEEPG--DIVAKIELPIRKNHIDLIDVFRKKRKNLIPRYVALIDREAKSIX 137  
288 ESMTEDEYSGPLDEGKVTISPMKE---LMQWYDQK-KLHRKCAVQLIVYKEXVL 341  
138 QPNTSPSTAVSQVTVCGDLHGKLDLILVYLHKKGLPSSSNPY----- 183  
342 KLSIVETLTKETKLTTCGDTHTGQFYDLNIFELNGLPSEETNPVYSPSAPPSPHLHPQ 401  
184 -----FNGDFVDRGKGLVLLLSLYLAF 209  
402 PPOVCTEWBEPPLPGKDVTLERRGLAMLHGQIGFDDVDKGSFSEVILTLFGKLYI 461  
210 PNAVPLNRGNHSDSVNNAARYGFLREYASKYPRNHRKILAFIDVYRWMLPLGSLNSRVLI 269  
462 PDHFLRLRGNHETDNNQIYGFGEVYKAYT---AQMYELPSFVEFWMLACINGKXLI 518  
270 VNGGF--SDTSLDLIKSIDRGKYVSIAPPLTDGSPPLDKTEMQOIFDIMSDPQ----- 322  
519 MGGGLPSEBDGVTLDTRKIERNR---QPP--DSALPRRMGFVPLARNSGSEPPPHLR 571  
323 -----ATMGCVNTL-----RGAGVAFGSDVYDNLQ 349  
572 LDPVHLHGAVMASVTVTWVRGEGVTAHTCPYLFNGSISIKRGVSCFGSDVYTAFLR 631  
350 RRLSLVIRSHCKPQGHFPMHDKIITIFASNYAIGSNKGAAYRL--NQLMPHFVY 408  
632 ENNLVDIIRSHKAGVAGVAGRGCVTYFSAFNVYCDQGNKASYYHLQGSDLRPFQHP 691  
RESULT 20  
US-10-369-493-21980

; Sequence 21980, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 21980  
; LENGTH: 513  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae  
US-10-369-493-21980  
Query Match 13.9%; Score 479.5; DB 4; Length 513;  
Best Local Similarity 32.3%; Pred. No. 1.2e-30;  
Matches 130; Conservative 68; Mismatches 154; Indels 51; Gaps 13;  
2 DENAIRAIFIQWRRHQARREM---QRRCNQIFQNTLEYASBDQAEIKYFFNDLIK 58  
113 DPATKALTCDRFREERFRKATIGAGNEAKISLCOTLNUISFPDANLDANY----- 165  
59 MPQAGRNQYQGSANHVSLDDKDLVEFGDIVAKIELPIRKNHIDLIDVFRKKRN 118  
166 ---RGPLERF-----QLYDDKN-----AFKAKIKMNSQEFISGVNDLFLK--GK 207  
119 RLHPRTVALILREAKSINKULPNISPV--STAVSQVTVCGDLHGKLDLVLVLRKGLP 176  
208 YLPKTYVALIISHADTLFRQEPSWVLENNSTPVDKISVCCDTHQGFYDVNLFRKFGV 267  
177 SSSNRYVNGPVPVRGKRGLEVLILLSLYLAFPAVAVLNGNHSDVMNARYGIRRYE 236  
268 GKPTHTLFGDPVDGSSCEVALLFYCLKILHPNPLFNGNHSDNNKIKYGEDECK 327  
237 SKYPRNHRKILAFIDVYRWMLPLGSLNSRVLIYHGF--SD--STSLDLIKSIDRGKYVI 294  
328 YKY---SGRIFNMFAQSFSPLATLINDVIVHGLPSDSATLSDFKNIDR-----F 379  
295 LRPPLTDGSPPLDKTEMQOIFDIMSDPQATMGCVNTLRGAGVWFGPVTNDFLQRHRS 354  
380 AQP--RDG-----AFMELTMADPOBANGMGPSQ--RGLGHAFGPDIITRFFLRNKL 428  
355 VYIRSHCKPQGHFPMHDKIITIFASNYAIGSNKGAAYRL 397  
429 KLFPSHRLMGVGPQKQKLMVTVASPNYCDQSGNLGVLHV 471  
RESULT 21  
US-11-097-143-5988  
; Sequence 5988, Application US/11097143  
; Publication No. US20050208558A1  
; GENERAL INFORMATION:  
; APPLICANT: Venter, J. Craig  
; APPLICANT: et al.  
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
; FILE REFERENCE: CLO00728  
; CURRENT APPLICATION NUMBER: US/11/097,143  
; CURRENT FILING DATE: 2005-04-04  
; PRIOR APPLICATION NUMBER: 60/157,832  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: 60/160,191  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR APPLICATION NUMBER: 60/161,932









Best Local Similarity 30.6%; Pred. No. 7.5e-27;  
Matches 113; Conservative 67; Mismatches 134; Indels 55; Gaps 13;

QY 60 POAAGKNOYQGSAAHVSVDKDDLYBEBRDYNAK--IELPRKNHIDLLIDVPRKKG 117  
DB 16 PAAGG---QGGGIDALL---DDIRRLLEVTAARPGKQVQVLSBSIRQLCTVSR---- 65  
QY 118 NRLHPKVALILREBAKSLKQLPNISPVSTAASQVTVCGDLHGKLDLLVLAHKGGLPS 177  
DB 66 -----AIFLSQ-----PNLELEAPIK-----ICGDHIGQVSDLLRLEFYGGFPP 105  
QY 178 SSNPVFNQDFVDRGKRGLEVLILLLSLYAFPAVFLNKGHEDSVNARVGFIREVS 237  
DB 106 EAN-YLFLGDYVDRGKQLETTICLLAYKIKYKYPENFLLRGNECASINRIYGFDECKR 164  
QY 238 KYPRNKRILAFIDYVRLPLGSVLSNRLVIHGGFS-DSTSLDLIKSIDRGKYSILR 296  
DB 165 RF---NVLRMKVTFECFNTLPVALIIDDKILCMHGLSPDLAHLDEIKMLQR----- 213  
QY 297 PPLTDEBPLDKTEWQOIFDIMSDP-QATWGCVPNTLRGAGVFGPDVTDNFLQRRRLSY 355  
DB 214 -----PTDVPGGLCDLWSDPKQVQMGAND-RGVSYTFGADKYSRFLQRRHDL 265  
QY 356 VIRSHCKENGHEFMHDKITITIFSAANTYALGSKNGAYIRLNQMLPHFVQYISAASOT 415  
DB 266 ICRAHQVEDGYEFPAADRLVITIFSAPNYCGEFDNAGAMMSVDETLMCSF-QLKPAERK 324  
QY 416 KRLSPKQRM 424  
DB 325 QIYGAKQNV 333

## RESULT 30

US-10-425-114-70117  
; Sequence 70117, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jindong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven B  
; APPLICANT: Tabaska, Jack B  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 70117  
; LENGTH: 334  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17236G06\_FTL.pep  
US-10-425-114-70117

Query Match 12.5%; Score 430.5; DB 4; Length 334;  
Best Local Similarity 30.6%; Pred. No. 7.5e-27;

Matches 113; Conservative 67; Mismatches 134; Indels 55; Gaps 13;

QY 60 POAAGKNOYQGSAAHVSVDKDDLYBEBRDYNAK--IELPRKNHIDLLIDVPRKKG 117  
DB 16 PAAGG---QGGGIDALL---DDIRRLLEVTAARPGKQVQVLSBSIRQLCTVSR---- 65  
QY 118 NRLHPKVALILREBAKSLKQLPNISPVSTAASQVTVCGDLHGKLDLLVLAHKGGLPS 177  
DB 66 -----AIFLSQ-----PNLELEAPIK-----ICGDHIGQVSDLLRLEFYGGFPP 105  
QY 178 SSNPVFNQDFVDRGKRGLEVLILLLSLYAFPAVFLNKGHEDSVNARVGFIREVS 237  
DB 106 EAN-YLFLGDYVDRGKQLETTICLLAYKIKYKYPENFLLRGNECASINRIYGFDECKR 164

QY 238 KYPRNKRILAFIDYVRLPLGSVLSNRLVIHGGFS-DSTSLDLIKSIDRGKYSILR 296  
DB 165 RF---NVLRMKVTFECFNTLPVALIIDDKILCMHGLSPDLAHLDEIKMLQR----- 213

QY 297 PPLTDEBPLDKTEWQOIFDIMSDP-QATWGCVPNTLRGAGVFGPDVTDNFLQRRRLSY 355  
DB 214 -----PTDVPGGLCDLWSDPKQVQMGAND-RGVSYTFGADKYSRFLQRRHDL 265

QY 356 VIRSHCKENGHEFMHDKITITIFSAANTYALGSKNGAYIRLNQMLPHFVQYISAASOT 415  
DB 266 ICRAHQVEDGYEFPAADRLVITIFSAPNYCGEFDNAGAMMSVDETLMCSF-QLKPAERK 324

QY 416 KRLSPKQRM 424  
DB 325 QIYGAKQNV 333

## RESULT 31

US-10-739-930-7415  
; Sequence 7415, Application US/10739930  
; Publication No. US20040216190A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH  
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT  
; FILE REFERENCE: 38-21(53377)B  
; CURRENT APPLICATION NUMBER: US/10/739,930  
; CURRENT FILING DATE: 2003-12-18  
; NUMBER OF SEQ ID NOS: 11088  
; SEQ ID NO 7415  
; LENGTH: 325  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: ZEMA-23APR03-C14584\_1.P  
US-10-739-930-7415

Query Match 12.5%; Score 429.5; DB 5; Length 325;  
Best Local Similarity 31.6%; Pred. No. 8.7e-27;

Matches 99; Conservative 67; Mismatches 118; Indels 29; Gaps 6;

QY 94 AKIEPRKNHIDLLIDVPRKRGKRNRLHPKYVALILREBAKSLKQLPNISPVSTAASQV 153  
DB 14 AAVDEVVRR-----LYBGGRRGRQVQMSABIRQLCVAKQVLLSQPNTL-----LRTHAPV 64

QY 154 TVCGDLHGKLDLLVLAHKGGLPSNPNVFNQDFVDRGKRGLEVLILLLSLYAFPAV 213  
DB 65 KICGDHIGQFVDLRLFLDYGYPPTST-YIFLGDYVDRGKQLETTICLLAYKGLKYPDI 123

QY 214 FLNRGNHEDSVNARVGFIREVSKYPRNKRILAFIDYVRLPLGSVLSNRLVIHGG 273  
DB 124 YLLRGNEHDAKINRYVGFDECKRRF--NVLKMKIFCCCFNCLPMAALIIDDKILCMHG 180

QY 274 FS-DSTSLDLIKSIDRGKYSILRPELTGEBPLDKTEWQOIFDIMSDPQATMGCVPTL 332  
DB 181 LSPBELSLDQIDKER-----PTRIDYGLCDLWSDPSDHTBGEGBSD 225

QY 333 RGAGVFGPDVTDNFLQRRRLSYVIRSHCKENGHEFMHDKITITIFSAANTYALGSKNG 392  
DB 226 RGVSCFFGADKVEFLKNDLVLVCRAHQVVDGYEFPAERRLVITIFSAPNYCGEFDNAG 285

QY 393 AVIRLNQMLPHF 405  
DB 286 ALISIDESIMCSF 298

## RESULT 32

US-10-739-930-9330  
; Sequence 9330, Application US/10739930  
; Publication No. US20040216190A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH

;; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT  
;; FILE REFERENCE: 38-21(53377)B  
;; CURRENT APPLICATION NUMBER: US/10/739,930  
;; CURRENT FILING DATE: 2003-12-18  
;; NUMBER OF SEQ ID NOS: 11088  
;; SEQ ID NO 9330  
;; LENGTH: 326  
;; TYPE: PRT  
;; ORGANISM: Glycine max  
;; FEATURE:  
;; OTHER INFORMATION: Clone ID: GLYMA-23APR03-C7096\_15.p  
US-10-739-930-9330

Query Match 12.5%; Score 429; DB 5; Length 326;  
Best Local Similarity 30.5%; Pred. No. 8; 7e-27;  
Matches 102; Conservative 66; Mismatches 125; Indels 41; Gaps 8;

QY 90 DIVNAKLEIPRKNHIDLIDVPRKRGKGN--RLHPKVALILREAAKSLKQLPNISPVST 147  
DB 8 DIINRLLEV-----RGRGKQVQLSESIKRLCAASREIFLQENLLELA 53  
QY 148 AVSQOVTVCGLDHLGKLDLLVLHKNGLPSSSNPYVNGDFVDRGKRGLEVLILLLSYL 207  
DB 54 PIK---ICGDVHGQYSDILRLFEYGGLPPEAN-YLFLGDYVDRGKQSLFTICLLAYKI 108  
QY 208 APPNAVFLNRGNHEDSVNNARYGFIREVESKYPRNKHILAFIDVYRMLPLGSVANSRV 267  
DB 109 KYPENFLLRGNHRCASINRIYGFDECKRRF--NVRIMKPTCEFCNCLPVALILDEKI 165  
QY 268 LIVHGGSF-DSTSLDLIKSIDRGKVVSIARPLTDGBPLDKTEMQOIFDIMWSDPQATWG 326  
DB 166 LCHHGSLSPDLMLDQIRNIR-----PVDVPTGGLCDLMLSDPSKEVQ 210  
QY 327 CVENTLRGAGWGPVDTNPLQRRRLSYVIRSHCKPNHGFMDNKKIITIFSASNYTA 386  
DB 211 GGMGNDRGVSYTFGADKVSFLLQKHDLILCRAHQVVEDGYEFAPARQLVTITISAPYCG 270  
QY 387 IGSNKAYIRLNQMLMPHFQYISAAQSQTRLSF 420  
DB 271 BFDNAGAMMSVDETLMCSF-QILKPADKRVKLN 303

RESULT 33  
US-10-425-114-62750  
;; Sequence 62750, Application US/10425114  
;; Publication No. US20040034888A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Zhou, Yihong  
;; APPLICANT: Kovalic, David K.  
;; APPLICANT: Screen, Steven E  
;; APPLICANT: Tabaska, Jack E  
;; APPLICANT: Cao, Yongwei  
;; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
;; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
;; FILE REFERENCE: 38-21(53313)B  
;; CURRENT APPLICATION NUMBER: US/10/425,114  
;; CURRENT FILING DATE: 2003-04-28  
;; NUMBER OF SEQ ID NOS: 73128  
;; SEQ ID NO 62750  
;; LENGTH: 338  
;; TYPE: PRT  
;; ORGANISM: Zea mays  
;; FEATURE:  
;; OTHER INFORMATION: Clone ID: LIB3357-018-H5\_P1.p  
US-10-425-114-62750

Query Match 12.5%; Score 429; DB 4; Length 338;  
Best Local Similarity 31.8%; Pred. No. 1e-26;  
Matches 103; Conservative 60; Mismatches 129; Indels 32; Gaps 7;

QY 94 AKIELEPIR-----KNHIDLIDVPRKRGKGNRLHPKYVALI---LREAAKSLKQLPNI 142

DB 1 AGISLPSRPFPGANDBAVDDLIRRLERAGRT--PRNAQVDAIRRLCAAAKOVFLS 59  
QY 143 SPVSTAVSQOVTVCGLDHLGKLDLLVLHKNGLPSSSNPYVNGDFVDRGKRGLEVLILL 202  
DB 60 QPILLELAPIKICGVHGYSDILRLFEYGGYPPDAN-YLFLGDYVDRGKQSLFTICLL 118  
QY 203 LSLYLAFPNNAVFLNRGNHEDSVNNARYGFIREVESKYPRNKHILAFIDVYRMLPLGSV 262  
DB 119 LAYKIRYENFLLRGNHRCASINRIYGFDECKRRF--NVRIMKPTCEFCNCLPVAL 175  
QY 263 LNSRVLIVHGGSF-DSTSLDLIKSIDRGKVVSIARPLTDGBPLDKTEMQOIFDIMWSDP 321  
DB 176 IDDKIFCHHGSLSPDLKSMQIRNIR-----PVDVPTGGLCDLMLSDP 220  
QY 322 QATMGCVPTLRGAGWGPVDTNPLQRRRLSYVIRSHCKPNHGFMDNKKIITIFS 381  
DB 221 DKXIDRWGNDRGVSYTFGADVAAEFLQKHDLILCRAHQVVEDGYEFAPKQLVITIFSA 280  
QY 382 SNYTAIGSNKAYIRLNQMLMPHF 405  
DB 281 PNICGEFDNAGALMSIDNSLVCSE 304

RESULT 34  
US-10-425-114-66618  
;; Sequence 66618, Application US/10425114  
;; Publication No. US20040034888A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Zhou, Yihong  
;; APPLICANT: Kovalic, David K.  
;; APPLICANT: Screen, Steven E  
;; APPLICANT: Tabaska, Jack E  
;; APPLICANT: Cao, Yongwei  
;; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
;; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
;; FILE REFERENCE: 38-21(53313)B  
;; CURRENT APPLICATION NUMBER: US/10/425,114  
;; CURRENT FILING DATE: 2003-04-28  
;; NUMBER OF SEQ ID NOS: 73128  
;; SEQ ID NO 66618  
;; LENGTH: 338  
;; TYPE: PRT  
;; ORGANISM: Zea mays  
;; FEATURE:  
;; OTHER INFORMATION: Clone ID: LIB4740-047-C1\_P1.p  
US-10-425-114-66618

Query Match 12.5%; Score 429; DB 4; Length 338;  
Best Local Similarity 31.8%; Pred. No. 1e-26;  
Matches 103; Conservative 60; Mismatches 129; Indels 32; Gaps 7;

QY 94 AKIELEPIR-----KNHIDLIDVPRKRGKGNRLHPKYVALI---LREAAKSLKQLPNI 142  
DB 1 AGISLPSRPFPGANDBAVDDLIRRLERAGRT--PRNAQVDAIRRLCAAAKOVFLS 59  
QY 143 SPVSTAVSQOVTVCGLDHLGKLDLLVLHKNGLPSSSNPYVNGDFVDRGKRGLEVLILL 202  
DB 60 QPILLELAPIKICGVHGYSDILRLFEYGGYPPDAN-YLFLGDYVDRGKQSLFTICLL 118  
QY 203 LSLYLAFPNNAVFLNRGNHEDSVNNARYGFIREVESKYPRNKHILAFIDVYRMLPLGSV 262  
DB 119 LAYKIRYENFLLRGNHRCASINRIYGFDECKRRF--NVRIMKPTCEFCNCLPVAL 175  
QY 263 LNSRVLIVHGGSF-DSTSLDLIKSIDRGKVVSIARPLTDGBPLDKTEMQOIFDIMWSDP 321  
DB 176 IDDKIFCHHGSLSPDLKSMQIRNIR-----PVDVPTGGLCDLMLSDP 220  
QY 322 QATMGCVPTLRGAGWGPVDTNPLQRRRLSYVIRSHCKPNHGFMDNKKIITIFS 381  
DB 221 DKXIDRWGNDRGVSYTFGADVAAEFLQKHDLILCRAHQVVEDGYEFAPKQLVITIFSA 280  
QY 382 SNYTAIGSNKAYIRLNQMLMPHF 405



Matches 112; Conservative 67; Mismatches 135; Indels 55; Gaps 13;

QY 60 POAGRKNOYQSAAHVSVDLDDKDDVBERGDIYNAK--ISLPKRNHIDLLIDVPRKRG 117  
DB 5 PAAG---GGGGGIDALL---DDIRLLLEVTRARQVQLSSESRQLCTVSR----54  
QY 118 NRLHPKVALLIREAAKSLKQLPNISPVSPAQVAVCGDLHAKGLDILLVHLKNGLPS 177  
DB 55 -----AIFLQ-----PNLELEAPIK-----ICDHHQVSDLRLEBYGGFP 94  
QY 178 SSNPVYFNGDFVDRGKRGLEVLNLLSLYAPPAVFLNKGHDSVMNARYGPIREVS 237  
DB 95 EAN-YLFLGDIYVDRGKQSLFTICLLAYKIKYKYPENFLKGNHCASINRIYGFYDECKR 153  
QY 238 KYPRNKRILAFIDVYRMLPLGSLVNSRVLIVHGFS-DSTSLDIKSIDRGKYVSIIR 236  
DB 154 RF---NVRLMKVTECFEFTLPVALIDDKILCHGGISPLAHIDRDKNQR-----202  
QY 297 PPLTDGEPDLKTEMQOIFDIWMSDP-QATWGCVPNTLRGAGVGPVTDNPLQRRLSY 355  
DB 203 -----PFDVPOGLICDLMSDPGKDVQSGKMD-RGVSYTFGADKVEFLQRHDL 254  
QY 356 VIRSHCKPENGHEFMHDKIITIFSAANYAIGSNKGAAYIRLNQMLPHFVQYISAASOT 415  
DB 255 ICAHQVVEDGYEFEPDRQLVTIFSAPNYCGEFDNAGMMSVDETLKCSF-QILKPAERK 313  
QY 416 KRLSPKORM 424  
DB 314 XKFGPNKM 322

RESULT 38  
US-10-437-963-111414  
; Sequence 111414, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 111414  
; LENGTH: 322  
; TYPE: PRP  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_15395C.1.pep  
US-10-437-963-111414

Query Match 12.4%; Score 427; DB 4; Length 322;  
Best Local Similarity 30.4%; Pred. No. 1.4e-26;  
Matches 114; Conservative 63; Mismatches 136; Indels 62; Gaps 12;

QY 59 MPQAKRNOYQSAHVSVDLDDKDDVBERGDIYNAKIELPIRKNHIDLLIDVPRKRG 118  
DB 1 MAAPAGAGGGGGMAVLLDD-----IIR-----LLEVRTARPG-36  
QY 119 RLHPKVALLIREAAKSLKQLPNIS-----PVSTAVSQVAVCGDLHAKGLDILLVHL 171  
DB 37 ---KOVQ---SESRQLCTVSRSEIFLSQPNLLEAPIKICGDHGOVSDLRLE 88  
QY 172 KNGLPSSNPVYFNGDFVDRGKRGLEVLNLLSLYAPPAVFLNKGHDSVMNARYG 231  
DB 89 YGFPPEAN-YLFLGDIYVDRGKQSLFTICLLAYKIKYKYPENFLKGNHCASINRIYGF 147

QY 232 IREVESKPRNKRILAFIDVYRMLPLGSLVNSRVLIVHGFS-DSTSLDIKSIDRGK 230  
DB 148 YDECKRF---NVRLMKVTECFEFTLPVALIDDKILCHGGISPLTHIDRDKSLPR--202  
QY 291 YVSIILAPPLTDGEPDLKTEMQOIFDIWMSDP-QATWGCVPNTLRGAGVGPVTDNPLQ 349  
DB 203 -----PFDVPOGLICDLMSDPGKDVQSGKMD-RGVSYTFGADKVEFLR 248  
QY 350 RRLSVIRSHCKPENGHEFMHDKIITIFSAANYAIGSNKGAAYIRLNQMLPHFVQYI 409  
DB 249 KHDLDLCAHQVVEDGYEFEPDRQLVTIFSAPNYCGEFDNAGMMSVDETLKCSF-QIL 307  
QY 410 SAASQTKRLSPKORM 424  
DB 308 KPAERKGFMAKNM 322

RESULT 39  
US-10-424-599-194787  
; Sequence 194787, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 194787  
; LENGTH: 329  
; TYPE: PRP  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_17920C.1.pep  
US-10-424-599-194787

Query Match 12.4%; Score 427; DB 4; Length 329;  
Best Local Similarity 31.1%; Pred. No. 1.4e-26;  
Matches 107; Conservative 68; Mismatches 127; Indels 42; Gaps 10;

QY 105 IDLLIDVPRKRNHIDLLIDVPRKRGLEVLNLLSLYAPPAVFLNRYGPIREVS 157  
DB 10 INRLLEV-RGRPGKQVQ-----LSEA--EIKQLCTVSRDIFLRQPNLLEAPIKIG 59  
QY 158 DLHAKGLDILLVHLKNGLPSSNPVYFNGDFVDRGKRGLEVLNLLSLYAPPAVFLNR 217  
DB 60 DVHGOVSDLRLEFEGGLPPRSN-YLFLGDIYVDRGKQSLFTICLLAYKIKYKYPENFLR 118  
QY 218 GHHDSVMNARYGPIREVESKPRNKRILAFIDVYRMLPLGSLVNSRVLIVHGFS-D 276  
DB 119 GHHDSVMNARYGPIREVESKPRNKRILAFIDVYRMLPLGSLVNSRVLIVHGFS-D 276  
QY 277 STSLDIKSIDRGKYVSIIRPPLTDGEPDLKTEMQOIFDIWMSDPQATWGCVPNTLRGAG 336  
DB 176 LHNQIKLPR-----PIEVETGLICDLMSDPSDIDRGGENERGVS 220  
QY 337 VWFQPDVTDNPLQRRLSVIRSHCKPENGHEFMHDKIITIFSAANYAIGSNKGAAYR 396  
DB 221 YTFGADRVTEFLQKHDLDLCAHQVVEDGYEFEPDRQLVTIFSAPNYCGEFDNAGAMT 280  
QY 397 LNNQMLPHFVQYISAASQTKRLSPKORMIVSSALKELAVMR 440  
DB 281 VDETLVCSF-QILKVENKRPSK---GFGSTTVKQSTTRAK 319

RESULT 40  
US-10-425-115-283336  
; Sequence 283336, Application US/10425115





GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 20, 2006, 19:52:48 ; Search time 32 Seconds

(without alignments)  
209,328 Million cell updates/sec

Title: US-09-463-733-1

Perfect score: 3442

Sequence: 1 MDENAIRAIFIGKWRBHQ.....VEHDIDPTDCKSKVIDPKKS 661

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 70606 seqs, 10133881 residues

Total number of hits satisfying chosen parameters: 70606

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: Published Applications AA.New:\*  
2: /cgn2\_6/prodata/1/pubppaa/US08\_NEW\_PUB.pep:\*  
3: /cgn2\_6/prodata/1/pubppaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/prodata/1/pubppaa/US07\_NEW\_PUB.pep:\*  
5: /cgn2\_6/prodata/1/pubppaa/PCT\_NEW\_PUB.pep:\*  
6: /cgn2\_6/prodata/1/pubppaa/US09\_NEW\_PUB.pep:\*  
7: /cgn2\_6/prodata/1/pubppaa/US10\_NEW\_PUB.pep:\*  
8: /cgn2\_6/prodata/1/pubppaa/US60\_NEW\_PUB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	405.5	11.8	319	7	US-11-109-156-38 Sequence 38, App1
2	405.5	11.8	323	6	US-10-878-556A-184 Sequence 184, App1
3	405.5	11.8	323	7	US-11-109-156-37 Sequence 37, App1
4	391.5	11.4	521	7	US-11-109-156-34 Sequence 34, App1
5	379	11.0	309	7	US-11-109-156-24 Sequence 24, App1
6	379	11.0	309	7	US-11-109-156-39 Sequence 39, App1
7	121	3.5	1614	6	US-10-821-234-903 Sequence 903, App
8	111	3.2	548	6	US-10-995-561-810 Sequence 810, App
9	111	3.2	567	6	US-10-995-561-813 Sequence 813, App
10	111	3.2	559	6	US-10-995-561-812 Sequence 812, App
11	108	3.1	709	7	US-11-074-176-158 Sequence 158, App
12	107.5	3.1	694	7	US-11-202-516-10 Sequence 10, App1
13	107.5	3.1	752	7	US-11-202-516-8 Sequence 8, App1
14	105	3.1	5034	6	US-10-793-626-2964 Sequence 2964, App
15	104.5	3.0	1070	7	US/11/062 Sequence 4, App1
16	104.5	3.0	1095	7	US/11/062 Sequence 7, App1
17	103	3.0	270	6	US-10-118-590-18 Sequence 18, App1
18	103	3.0	1765	6	US-10-055-877-140 Sequence 140, App
19	102	3.0	149	6	US-10-821-234-1535 Sequence 1535, App
20	102	3.0	149	7	US-11-124-368A-268 Sequence 268, App
21	102	3.0	149	7	US-11-124-368A-269 Sequence 269, App
22	102	3.0	171	6	US-10-821-234-935 Sequence 935, App1
23	101.5	2.9	148	7	US-11-065-943-46 Sequence 46, App1
24	101	2.9	257	6	US-10-118-590-16 Sequence 16, App1
25	100.5	2.9	1548	7	US-11-108-172-1095 Sequence 1095, App

26	99	2.9	276	6	US-10-467-657-2502 Sequence 2502, App
27	98.5	2.9	746	7	US-11-074-176-314 Sequence 314, App
28	98.5	2.9	749	7	US-11-074-176-54 Sequence 54, App1
29	98	2.8	1538	6	US-10-995-561-772 Sequence 772, App
30	98	2.8	5335	6	US-10-995-561-777 Sequence 777, App
31	98	2.8	5406	6	US-10-995-561-774 Sequence 774, App
32	98	2.8	5415	6	US-10-995-561-779 Sequence 779, App
33	98	2.8	5464	6	US-10-995-561-775 Sequence 775, App
34	98	2.8	5935	6	US-10-995-561-776 Sequence 776, App
35	97.5	2.8	467	7	US-11-156-084-115 Sequence 115, App
36	97.5	2.8	613	6	US-10-131-826A-190 Sequence 190, App
37	97	2.8	216	6	US-10-118-590-2 Sequence 2, App1
38	97	2.8	216	6	US-10-118-590-6 Sequence 6, App1
39	97	2.8	814	6	US-10-878-556A-161 Sequence 161, App
40	97	2.8	1184	7	US-11-115-639-50 Sequence 50, App1
41	97	2.8	4384	6	US-10-821-234-1120 Sequence 1120, App
42	96.5	2.8	1145	6	US-10-793-626-1432 Sequence 1432, App
43	95.5	2.8	2897	6	US-10-499-715-2 Sequence 2, App1
44	95	2.8	225	6	US-10-118-590-30 Sequence 30, App1
45	95	2.8	252	6	US-10-118-590-22 Sequence 22, App1

## ALIGNMENTS

RESULT 1  
US-11-109-156-38  
; Sequence 38, Application US/11/09156  
; Publication No. US20050250144A1  
; GENERAL INFORMATION:  
; APPLICANT: Toshio Oca  
; APPLICANT: Takao Isogai  
; APPLICANT: Tetsuo Nishikawa  
; APPLICANT: Koji Hayashi  
; APPLICANT: Kaoru Otsuka  
; APPLICANT: Jun-ichi Yamamoto  
; APPLICANT: Shizuko Ishii  
; APPLICANT: Tomoyasu Sugiyama  
; APPLICANT: Ai Wakamatsu  
; APPLICANT: Keiichi Nagai  
; APPLICANT: Tetsuji Otsuki  
; APPLICANT: Shin-ichi Funahashi  
; APPLICANT: Chiaki Senoo  
; APPLICANT: Jun-ichi Nezu  
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN  
; FILE REFERENCE: 06501-099002  
; CURRENT APPLICATION NUMBER: US/11/109,156  
; PRIOR FILING DATE: 2005-04-19  
; PRIOR APPLICATION NUMBER: US/10/060,065  
; PRIOR FILING DATE: 2002-01-29  
; PRIOR APPLICATION NUMBER: PCT/JP00/05061  
; PRIOR FILING DATE: 2000-07-28  
; PRIOR APPLICATION NUMBER: US 60/159,590  
; PRIOR FILING DATE: 1999-10-18  
; PRIOR APPLICATION NUMBER: US 60/183,322  
; PRIOR FILING DATE: 2000-02-17  
; PRIOR APPLICATION NUMBER: JP 11-248036  
; PRIOR FILING DATE: 1999-07-29  
; PRIOR APPLICATION NUMBER: JP 2000-118776  
; PRIOR FILING DATE: 2000-01-11  
; PRIOR APPLICATION NUMBER: JP 2000-183767  
; PRIOR FILING DATE: 2000-05-02  
; PRIOR APPLICATION NUMBER: JP 2000-241899  
; PRIOR FILING DATE: 2000-06-09  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 38  
; LENGTH: 319  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-109-156-38

Query Match	11.8%	Score 405.5	DB 7	Length 319
Best Local Similarity	32.4%	Pred. No. 2.7e-26		
Matches	105	Conservative	58	Mismatches 126
				Indels 35
				Gaps 11
QY	102	KNHIDLIDVFEKKGNR-----LHPKVY-ALITREAKSKLQLENISFSTAVSQQT	154	
Db	2	KLNDISITQRLLEFVAGSKRGKRVQLOENIRGLCLKSRITFSQ-----PILSLEAPLK	56	
QY	155	VGCDLHGKLDLILVVLHKNGLPSSSNPYFENGDFVDRGKRGLEVLTLISLTYLAPNAYF	214	
Db	57	ICGDHIGQYDYLRLFEYGGFPRESN-YFLGDYVDGRQSLTICLLAYKIKYENPF	115	
QY	215	LNRGHNSVMMARAGTIREVESKTPRNHNRLLAFIDEVRYRLPLGASVANSVLYHGGF	274	
Db	116	LIRGHBCASINRIKGFYDECKRRY-NIKMKLTLD-CENCLPIALVDEKIFCCHGL	172	
QY	275	S-DSTSLDLIKSIDRGKYSILRPLTDSRPLDKTSMQDIPDIMSDF-QATWGCVPTL	332	
Db	173	SPDLQSMQIRIKMR-----PDEVDPQGLICLLMSDPRKDYLGAGEND	216	
QY	333	RGAGVWFGPDVTNPLQRRHLSVYIRSHCECKNGHGFPMHDKIITTFASANYAIGSNKG	392	
Db	217	RGVSFTFGAEVYAKFLHKHIDLILICRAHQVEDVEYGFPAKRGQLVTLFSAPNYCGEFDNAG	276	
QY	393	AYIRLNQMLPHFVQYISAASQTK	416	
Db	277	AMMSVDETLMCSF-QILKPAEKCK	299	

```

RESULT 2 556A-184
US-10-878-556A-184
Sequence 184, Application US/10878556A
Publication No US2005026396A1
GENERAL INFORMATION:
APPLICANT: Hoffmann La-Roche Inc.
TITLE OF INVENTION: HCV regulated protein expression
FILE REFERENCE: 21762
CURRENT APPLICATION NUMBER: US/10/878,556A
CURRENT FILING DATE: 2004-06-28
NUMBER OF SEQ ID NOS: 199
SOFTWARE: PatentIn version 3.1
SEQ ID NO 184
LENGTH: 323
TYPE: PRF
ORGANISM: Homo sapiens
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: sw_hum/pp1q_human
DATABASE ENTRY DATE: 1994-06-01
US-10-878-556A-184

```

	Query Match	11.8%	Score 405.5	DB 6,	Length 323;
	Best Local Similarity	32.4%;	Pred. No.2.8e-24;		
	Matches	105;	Conservative	58;	Mismatches 126; Indels 35; Gaps 11
Qy	102	KNHIDLLDVFKEKKGNNR-----LHPKKV-ALILREAKSLPKOLPNISPVSTAVSQVT	154		
Dd	6	KLTNDITIGRLLEFVGSGSKRGKVQLQENIRGLCLKSRAIFLSQ-----PILLETAPLK	60		
Qy	155	VCGDLAGKLDDLVLVHKNGLFSSSNPYFENGDPVDGRGRLGVLLLLSLYLAAPNAV	214		
Dd	61	ICGDIHGYDDLRFLFEYGFPPENS-VLFDDYVDGRQSILETICLLAYIKKPENEF	119		
Qy	215	LNRGNHDSVMNARVGFIREVESKTPRNKRILAFIDEVRWLPLGSVINSRVLIYHG	274		
Dd	120	LLRGNHCASINRIYGVFYDECKRRY-NIKLMKTFTD-CFNCLPIAIVDBKIPCHGGL	176		
Qy	275	S-DSTSLDLIKSIDGKTVSILRPPLTDGEPLDKTEWQOIFIMMSDP-QATMGCVPTL	332		
Dd	177	SPDLQSMBOIRIRIMR-----PTDVPDQGLCDLLMSDPKDVLWGMEIND-	220		
Qy	333	RGAAGVFGGDYVDNFLQGRHLVSVIRSHCECKNGHHFMHDNKIITTFASNTYAIGSRG	392		
Dd	221	RGVSTTFAGAEVAKTLHRKIDLICAHQVVDVGGVEFFKROLVYTLFSAPNVCGERDANG	280		

Qy 393 AYIRLNNQIMPHFVQYISAASQTK 416  
| : : | : : |  
Db 281 AMMSVDETLKCSF-QILPKAEKK 303

```

RESULT 3
US-11-109-156-37
/ Sequence 37, Application US/11109156
/ Publication No. US20050250144A1
GENERAL INFORMATION:
/ APPLICANT: Toshio Ota
/ APPLICANT: Takao Isogai
/ APPLICANT: Tetsuo Nishikawa
/ APPLICANT: Koji Hayashi
/ APPLICANT: Kaoru Otsuka
/ APPLICANT: Jun-ichi Yamamoto
/ APPLICANT: Shizuko Ishii
/ APPLICANT: Tomoyasu Sugiyama
/ APPLICANT: Ai Wakamatsu
/ APPLICANT: Keiichi Nagai
/ APPLICANT: Tetsuji Otsuki
/ APPLICANT: Shin-ichi Punahashi
/ APPLICANT: Chiaki Senoo
/ APPLICANT: Jun-ichi Nezu
/ TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN
/ TITLE OF INVENTION: PHOSPHATASE
/ FILE REFERENCE: 06501-099002
/ CURRENT APPLICATION NUMBER: US/11/109,156
/ CURRENT FILING DATE: 2005-04-19
/ PRIOR APPLICATION NUMBER: US/10/060,065
/ PRIOR FILING DATE: 2002-01-29
/ PRIOR APPLICATION NUMBER: PCT/JP00/05061
/ PRIOR FILING DATE: 2000-07-28
/ PRIOR APPLICATION NUMBER: US 60/159,590
/ PRIOR FILING DATE: 1999-10-18
/ PRIOR APPLICATION NUMBER: US 60/183,322
/ PRIOR FILING DATE: 2000-02-17
/ PRIOR APPLICATION NUMBER: JP 11-248036
/ PRIOR FILING DATE: 1999-07-29
/ PRIOR APPLICATION NUMBER: JP 2000-118776
/ PRIOR FILING DATE: 2000-01-11
/ PRIOR APPLICATION NUMBER: JP 2000-183767
/ PRIOR FILING DATE: 2000-05-02
/ PRIOR APPLICATION NUMBER: JP 2000-241899
/ PRIOR FILING DATE: 2000-06-09
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 37
/ LENGTH: 323
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-11-109-156-37

```

	Query Match	11.8%	Score 405.5;	DB 7;	Length 323;
	Best Local Similarity	32.4%;	Pred. No. 2.8e-24;		
	Matches	105;	Conservative	58;	Mismatches 126; Indels 35; Gaps 11
Qy	102 KKHIDLIDVKKRGRNR-----LHPKCV-ALILREAAKLKOLPNISPVSTAVSQOVT	154			
Dd	KLTNDISIQRLLBVGSKRGKNVOLEWBEIRGLCLKSRIEFLSQ-----PILIEAPLK	60			
Qy	155 VCGDHLGDLDDLVLVHNGLPSSNPFVPFNGDFPRDKRGLEVLLLSLYLAFFNAVF	214			
Dd	61 ICGIHGGYDLLRFVEYGFPSPSN-VLFPSDYVDKQSLLETICLLAYKIKYPENVF	119			
Qy	215 LNRGHEDSVANARGFIREVESKPRPHKEILLAFIDEVSWMLPLGSIVRSVLIVHGCF	274			
Dd	120 LLRNHRCAASINRIYGFYEDECKRR-NIKMKITFTD-CFNCPRLPAIYVDKELPCCHGL	176			
Qy	275 S-DSTSLDLKSIDRGKYVSILRPPLTDGEPLDKTEMOQIFDIMNSDP-QATMGCVPTLT	332			
Dd	177 SPDIQSQQIARRIMK-----PLTVDDQGLLCDDLMSDDPKDVLTGMGEND-	220			

```
QY 333 RGAGWGPDPVTNDFLORHRLSYIRSHCEKPNHEPFMDNKKITITFSASNYTAISGNKG 392
DB 221 RGVSFTFGAENVAKFPLRKHDLICRAHQVVDGDEFPARQVLTLFSAFNYCGEFDNAG 280
QY 393 AYIRLNQNLMPHFVQYISAASQTK 416
DB 281 AMMSVDETLKCSF-QILKPAEKCK 303

RESULT 4
US-11-109-156-34
; Sequence 34, Application US/11109156
; Publication No. US20050250144A1
; GENERAL INFORMATION:
; APPLICANT: Toshio Ota
; APPLICANT: Takao Isegai
; APPLICANT: Tetsuo Nishikawa
; APPLICANT: Koji Hayashi
; APPLICANT: Kaoru Otsuka
; APPLICANT: Jun-ichi Yamamoto
; APPLICANT: Shizuko Ishii
; APPLICANT: Tomoyasu Sugiyama
; APPLICANT: Ai Makamatsu
; APPLICANT: Keiichi Nagai
; APPLICANT: Tetsuji Otsuki
; APPLICANT: Shin-ichi Funahashi
; APPLICANT: Jun-ichi Nezu
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN
; TITLE OF INVENTION: PHOSPHATASE
; FILE REFERENCE: 06501-099002
; CURRENT APPLICATION NUMBER: US/11/109,156
; CURRENT FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: US/10/060,065
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: PCT/JP00/05061
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: US 60/183,322
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: JP 11-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; LENGTH: 521
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-109-156-34

Query Match 11.4%; Score 391.5; DB 7; Length 521;
Best Local Similarity 31.1%; Pred. No. 6.8e-23; Indels 47; Gaps 10;
Matches 101; Conservative 56; Mismatches 119;

QY 102 KKHIDLLIDVFRKRGKRLHPKYVALILREBAKSLKQLPNISPVSTAVSQQVTVCGDLHG 161
DB 40 KPRVDILKAHMKR--GRLEBSVALRITREGASILRQEKULDIDA----PVTVCGBIHG 93
QY 162 KLDLVLVLAHKNGLPSSSNVYVRGDPVDRGKGLVLLLSLYLAFPNVAVPLRNGNR 221
DB 94 QFDFLMKLFEVGVGSPANTR-YLFLGDVYDRGYFSIECVLYLWMLKILYPTLFLRGNH 152
QY 222 DSVNNAARYGFIRESVSKYPRNHRKRIILAFIDVYVRMLPLGSLVNSRVLIYHGFS-DSTSL 280
DB 153 CRHLTEFTFKQCKCKIKY---SERVYDACMDAFDCFLAALMQQFLCVHGGLSPRINTL 209
```

```
QY 281 DLKSIDRCKRYYSILAPPLTDGEPPLDKTEWQOIFDTMSPDQATMG-----CVPTLR 333
DB 210 DDRIKDLRFK-----BPF-----AYGPMCDILMSDPLBDFGNKTOEHPTNVTVR 254
QY 334 GAGWFGPDPVTNDFLORHRLSYIRSHCEKPNHEPFMDNKKITITFSASNYTAI 387
DB 255 GCSYFTSYPAVCEFLQHNLLSLIRAHADADGIRMYRRSQTTGFPSSLITIFSAFNYLDV 314
QY 388 GSNKGAAYIRLNQLM-----PH 404
DB 315 YNNKAALVKYENNVMMIRQFNCSPH 339

RESULT 5
US-11-109-156-24
; Sequence 24, Application US/11109156
; Publication No. US20050250144A1
; GENERAL INFORMATION:
; APPLICANT: Toshio Ota
; APPLICANT: Takao Isegai
; APPLICANT: Tetsuo Nishikawa
; APPLICANT: Koji Hayashi
; APPLICANT: Kaoru Otsuka
; APPLICANT: Jun-ichi Yamamoto
; APPLICANT: Shizuko Ishii
; APPLICANT: Tomoyasu Sugiyama
; APPLICANT: Ai Makamatsu
; APPLICANT: Keiichi Nagai
; APPLICANT: Tetsuji Otsuki
; APPLICANT: Shin-ichi Funahashi
; APPLICANT: Jun-ichi Nezu
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN
; TITLE OF INVENTION: PHOSPHATASE
; FILE REFERENCE: 06501-099002
; CURRENT APPLICATION NUMBER: US/11/109,156
; CURRENT FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: US/10/060,065
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: PCT/JP00/05061
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: US 60/183,322
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: JP 11-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-109-156-24

Query Match 11.0%; Score 379; DB 7; Length 309;
Best Local Similarity 31.1%; Pred. No. 2.9e-22; Indels 28; Gaps 7;
Matches 91; Conservative 62; Mismatches 112;

QY 119 KILHPKYVALILREBAKSLKQLPNISPVSTAVSQQVTVCGDLHGKLDLVLVLAHKNGLPSS 178
DB 22 QLESQVKSILCEKAKKILTKESNVQEVRC---PVTVCGBIHGQFDFLMELFRIGKSPD 77
QY 179 SNPYVNGDPFVDRGKGLVLLLSLYLAFPNVAVPLRNGNHSVMAARYGFIRESVSK 238
DB 78 TN-YLPMGDIVDRGYSVETVTLVLAALKVRYRERTITILGNHSSROI TOYGYHDECLRK 136
```

Qy 239 YPNRKRLLAFIDEVVRMLPLGSSNSRLLVHGSGSS--TSLDIKSIDEGKAVSLRP 237  
 Db 137 Y--GNANVMKXFLTDFLYLPLTALVDQGI FCHGGSLSPIDLTDIRLDRLGVPHEGR 194  
 Qy 298 PLTDSGEPLDTIEMQQLFDIMSDP--QATMGCVPTLTAGAGVRRPDTNDFLORHLSY 355  
 Db 195 -----MCDLMSDPDRDGKQGLSP---RQAGTTFQGLDSLETFNANGILL 236  
 Qy 356 VTSHECKPENGHEFMDNKLITTFASANTYALIGSNKAYIRLNOIMPHVOY 408  
 Db 237 VSRPHQIWMGYNMCHDRVYITFSAIPNVCYRGNQOALIMELDITLKSPLOF 289

RESULT 6  
US-11-109-156-39

```

; Sequence 39, Application US/11109156
; Publication No. US20050250144A1
GENERAL INFORMATION:
APPLICANT: Toshio Ota
APPLICANT: Takao Isogai
APPLICANT: Tetsuo Nishikawa
APPLICANT: Koji Hayashi
APPLICANT: Kaoru Otsuka
APPLICANT: Jun-ichi Yamamoto
APPLICANT: Shizuko Ishii
APPLICANT: Tomoyasu Sugiyama
APPLICANT: Ai Wakamatsu
APPLICANT: Keiichi Nagai
APPLICANT: Tetsuji Otsuki
APPLICANT: Shin-ichi Funahashi
APPLICANT: Chiaki Senoo
APPLICANT: Jun-ichi Nezu
TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN
TITLE OF INVENTION: PHOSPHATASE
FILE REFERENCE: 06501-099002
CURRENT APPLICATION NUMBER: US/11/109,156
PRIORITY FILING DATE: 2005-04-19
PRIORITY APPLICATION NUMBER: US/10/060,065
PRIORITY FILING DATE: 2002-01-29
PRIORITY APPLICATION NUMBER: PCT/JP00/05061
PRIORITY FILING DATE: 2000-07-28
PRIORITY APPLICATION NUMBER: US 60/159,590
PRIORITY FILING DATE: 1999-10-18
PRIORITY APPLICATION NUMBER: US 60/183,322
PRIORITY FILING DATE: 2000-02-17
PRIORITY APPLICATION NUMBER: JP 11-248036
PRIORITY FILING DATE: 1999-07-29
PRIORITY APPLICATION NUMBER: JP 2000-118776
PRIORITY FILING DATE: 2000-01-11
PRIORITY APPLICATION NUMBER: JP 2000-183767
PRIORITY FILING DATE: 2000-05-02
PRIORITY APPLICATION NUMBER: JP 2000-241899
PRIORITY FILING DATE: 2000-06-09
PRIORITY APPLICATION NUMBER: JP 2000-06-09
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 39
LENGTH: 309
TYPE: PRT
ORGANISM: Homo sapiens
US-11-109-156-39

```

	Query Match	11.0%;	Score 379;	DB 7;	Length 309;
	Best Local Similarity	31.1%;	Pred.No.2.9e-22;		
	Matches	91;	Conservative 62;	Mismatches 112;	Indels 28; Gaps 7
Gy	119	RHPKVVLLIIEAAKSLKQLPNISPVSTAVAGQVWYVCGLDHGKDLDLVYLHNGKLPS	178		
	:	: : : : :	: : : : : :	: : :	
Dd	22	QISESQVSLCEKAKILTKESNVGEVRC---- <p>PVTVCADVHGQFHDMLERIGKSPD</p>	77		
Gy	179	SNPVFNGDFVGGRKGLEVLIIISLYLAFNVAFLNMGHEDSVMMARVGFIREVBK	238		
	:	: : : : : :	: : : : : :	: : :	
Dd	78	IN-YLFMGVYDVGRGYSAVELTTLVALAKRYEERTITLGNHSRQITGVGFYDECLRK	136		

Oy 239 YPNHHKRIIAFIDEVYRMFLGVSZSNVYLIVHGFGPS-TSLDLSISIDPKGVSTLR 237  
 Db 137 Y--GNANWKYFLFDLFDLYFLPALVDGQIFCHGSLSPSIDTLHIFALDRLGCVPHGR 194  
 Oy 298 FLIDGEPDLKTEMOQIFDIIMSDP--QATMGCVPTLPGAVMGCPVTNLFLORHLSY 355  
 Db 195 -----NCDLIMSDPDDRGKGISP---RGAITYTGODISFTFNHANGTL 236  
 Oy 356 VIRSHCKPENGHEFMDNKILITIFASNTYALGSKGAYIRLNNQIMPHVOY 408  
 Db 237 VSAHOLVMBGVYMWCDRAVVITIFASPNVCYRCGQAALIMELDITLKYSFQF 289

RESULT 7  
US-10-821-234-903

```

: Sequence 903, Application US/10821234
: Publication No. US20050255114A1
: GENERAL INFORMATION:
: APPLICANT: Labat, Ivan
: APPLICANT: Strache-Craign, Birgit
: APPLICANT: Andarmant, Susan
: APPLICANT: Tang, Y. Tom
: TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
: FILE REFERENCES: 821A
: CURRENT APPLICATION NUMBER: US/10/821,234
: CURRENT FILING DATE: 2004-04-07
: PRIOR APPLICATION NUMBER: US 60/462,047
: PRIOR FILING DATE: 2003-04-07
: NUMBER OF SEQ ID NOS: 1704
: SOFTWARE: pt_seq_genes Version 1.0
: SEQ ID NO 903
: LENGTH: 1614
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-10-821-234-903

```

Query Match	3.5%;	Score 121;	DB 6;	Length 1614;
Best Local Similarity	19.8%;	Pred. No. 0.33;		
Matches 168;	Conservative 122;	Mismatches 274;	Indels 284;	Gaps 47

QY	8	AAIFLOKRYRHOAKREORRCMNQIFONLEBAEOOALYKCFENDIKMPAOAGKN	67
Db	762	SVATIGSFMKATARKSYLSR-----LQFFDHNN-ELVK-----IGSLRANKARD	807
QY	68	QYQ-----GSAN-----VSUIDKCD-DLVEBFG-----DIV	92
Db	808	DYKTLVSENPPLTVIRKFFVYLLDQSDIDFQEBLEVARLREBVTKIRANQOLEKIDMLM	867
QY	93	NAKLELPKRNKHILLIDV-----FRKRG-----NRLHPKVALLIREAAKSILQLP	140
Db	868	DIKIGLVL-KNRI-TLEDVISHSKKLNKGGKGEHEILLNTDNOGKSISKRRKTLLEYQ	925
QY	141	NI-----SEPVSTAVSGOQVTVCGDLHG-----KLDDLLVVLHKNGLPS-	177
Db	926	QLFVILQTNPLYLAKLIFQMPQONSTKFMQDVIIFLVLYVANSNGQEBYLLLLFTYALBEE	985
QY	178	-----GSNPVY-----FNQDFUDRGKRGLEVLILLIS-----	204
Db	986	IKSKVDQVODIVTNPFPVIRKVVSEFN-----RGRAGONTLQLLAPVYKEIIDDKSLIN	1040
QY	205	-----LYLAEPNAVAVLNRGNHRDSVMNARYGFIREVSKYPRNKHRIALAFIDRYVR	255
Db	1041	TNPVEVTVKAMNOLETORG---EASKLPIVYTTBQALTYIEVANKLEASJENLRVYDK	1096
QY	256	-----WLPLGSAVLSREVL---IVHGEFSDBTSLDILKSIDRGKVSILRPPLT--D	301
Db	1097	VIANGISLDDLPLPGRLAIYAKVLKNSIHEKFPDATEDBLKTIQNLLYYRYMANPAIVAPD	1156
QY	302	G-BPLDKTEMOQOIFDIMSDDPQATMGCCPNTLRAG---VWFGSD-----VTDNFL-----Q	349
Db	1157	GFDIIDMTAGQOIN---SDQRNNGISYAKVLOHAAASKNLEEGENSEHLSMNWNLSETYQ	1212

```

QY 350 RHRLSVIRSHCK-PNGHFMHDK---ITTFASANYAIGSNKGAIVRLNOLMPH- 404
DB 1213 ERKTF--KACANPEPEKKNMDKYTLTV-----SKVYIISIBIISTHS 1259
QY 405 -FVOYISAQTRLSFKQRMGIV-----ESSALKEIAVRMD-----HRDELD 448
DB 1260 LLEHODALAPKNDLSELGLSVPTVESFLGAVDPNDPNKANTLSQLSKREISL 1319
QY 449 EF-RKYDPDQSGYISISHWCVMENYTKLGLPRLRLDKLAPCTDSQKVNRTLDLT 507
DB 1320 VLTSTKDEJEDGALISR---SIMGTKKLIID--VIRNQ--PG-----MTLFEILRT 1364
QY 508 DVLLEADGMSVMDALYANKASLVAFNIIADNNGEITLDEFEAIDLIV-AHNRGAV 566
DB 1365 PTAQOEVRHATDM---VSRA-----MIDSR-----TPBKRHSQSMIEDOULPLBO 1408
QY 567 SKAEMLEKCRMDLNG---DGKVD--LNEPLAFLSDLRKEODE----- 608
DB 1409 KKRKIQRNRLTEQCHVSEKKYODILNEIAKDIPNQRIRYKRAELAKLQOTLAINL 1468
QY 609 -----NIRRRST-----GRPSVAKTATDPVTLLADKISK 637
DB 1469 KKAAPYEQIINYDYIKTKLNDLKRKNTRRSIKLDGKGPYKAKRA-KPVYTYAKLHE 1527
QY 638 NTLVVEHD 645
DB 1528 KGVLLDID 1535

```

RESULT 8  
US-10-995-561-810  
Sequence 810, Application US/10995561  
Publication No. US20050272054A1

GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele et al.  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE. METHODS OF  
TITLE OF INVENTION: DETECTION AND USES THEREOF  
FILE REFERENCE: CL001559  
CURRENT APPLICATION NUMBER: US/10/995,561  
CURRENT FILING DATE: 2004-11-24  
NUMBER OF SEQ ID NOS: 85702  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 810  
LENGTH: 548  
TYPE: PRF  
ORGANISM: Homo sapiens  
US-10-995-561-810

Query Match 3.2%; Score 111; DB 6; Length 548;  
Best Local Similarity 18.6%; Pred. No. 0.4;  
Matches 86; Conservative 67; Mismatches 175; Indels 134; Gaps 19;

```

QY 273 GSDSTSLDLSIRGKVSILRPP--LTDGEPDKTWMQOIFDIMSDDPQATMGCYVN 330
DB 91 GPS-SFARSLVPSSE-LHLSLKSSKVITNGPEKNSSSPSSVDYAAAGPRKL----- 142
QY 331 TLRGAGVWGPVDTNLFQHRHLSYVIRSHCKPNG-----HEFMHDKII----- 376
DB 143 ---SSGALYGRPRSTFPNSHRFQLVISK--APSGDLIDKXSELFSNNQLPPTPTLKT 196
QY 377 ---TFASANYAIGSNKGAIVRLNQLMPHFVQYISAASQTRLSFKQRMGIVESSALK 433
DB 197 EAKSFLSQYRYTYTPAKRKDPFD-----QRIEATQTE-LSFKSELGTAFRTKMT 245
QY 434 BLAVRMRDR-----DELEDEFKRY----- 453
DB 246 DSEBNIKQANSNCVTYAKKIAPLPEGHSTWDEIKDQLQHSPPRMCQYSLKPPSTR 305
QY 454 ---DPKDSGYISISHWCVMENYTKLGL-----PWRLRLDKLAPCTDSQKVNRTLDL 504
DB 306 KIYSDBEELLYSFTB--DVTDEILKGLFSNRFLERLPERHIIKQKHLEBKMRHLHY 363

```

```

QY 505 LDTVILAEADGMSVMDALYANKASLVAFNIIADNNGEITLDEFEAIDLIVAMRG 564
DB 364 LNVDTAGCTSEENSVQND-----VDMLVNFPDEKAGSEBPELNKSEVTTIQOERO 414
QY 565 AYSKAEMLKCRMDLNGDKVDLNE-----PLEAFLSDLRH--KEODENIRRR 613
DB 415 QYOKA-----LMLASAPDENBIEIPSPTEFPMPYIKSHGSECVIIQVNDFTNLT 466
QY 614 ST---GRPSVAKTATDPVTLLADKISKNTLVVEHIDPTDCE 652
DB 467 STLDENHPSISDILDTRET-----SVNVIEGSDPEKVE 500

```

RESULT 9  
US-10-995-561-813  
Sequence 813, Application US/10995561  
Publication No. US20050272054A1  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele et al.  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE. METHODS OF  
TITLE OF INVENTION: DETECTION AND USES THEREOF  
FILE REFERENCE: CL001559  
CURRENT APPLICATION NUMBER: US/10/995,561  
CURRENT FILING DATE: 2004-11-24  
NUMBER OF SEQ ID NOS: 85702  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 813  
LENGTH: 567  
TYPE: PRF  
ORGANISM: Homo sapiens  
US-10-995-561-813

Query Match 3.2%; Score 111; DB 6; Length 567;  
Best Local Similarity 18.6%; Pred. No. 0.42;  
Matches 86; Conservative 67; Mismatches 175; Indels 134; Gaps 19;

```

QY 273 GSDSTSLDLSIRGKVSILRPP--LTDGEPDKTWMQOIFDIMSDDPQATMGCYVN 330
DB 110 GPS-SFARSLVPSSE-LHLSLKSSKVITNGPEKNSSSPSSVDYAAAGPRKL----- 161
QY 331 TLRGAGVWGPVDTNLFQHRHLSYVIRSHCKPNG-----HEFMHDKII----- 376
DB 162 ---SSGALYGRPRSTFPNSHRFQLVISK--APSGDLIDKXSELFSNNQLPPTPTLKT 215
QY 377 ---TFASANYAIGSNKGAIVRLNQLMPHFVQYISAASQTRLSFKQRMGIVESSALK 433
DB 216 EAKSFLSQYRYTYTPAKRKDPFD-----QRIEATQTE-LSFKSELGTAFRTKMT 264
QY 434 BLAVRMRDR-----DELEDEFKRY----- 453
DB 265 DSEBNIKQANSNCVTYADAKEIAPLPEGHSTWDEIKDQLQHSPPRMCQYSLKPPSTR 324
QY 454 ---DPKDSGYISISHWCVMENYTKLGL-----PWRLRLDKLAPCTDSQKVNRTLDL 504
DB 325 KIYSDBEELLYSFTB--DVTDEILKGLFSNRFLERLPERHIIKQKHLEBKMRHLHY 382
QY 505 LDTVILAEADGMSVMDALYANKASLVAFNIIADNNGEITLDEFEAIDLIVAMRG 564
DB 383 LNVDTAGCTSEENSVQND-----VDMLVNFPDEKAGSEBPELNKSEVTTIQOERO 433
QY 565 AYSKAEMLKCRMDLNGDKVDLNE-----PLEAFLSDLRH--KEODENIRRR 613
DB 434 QYOKA-----LMLASAPDENBIEIPSPTEFPMPYIKSHGSECVIIQVNDFTNLT 485
QY 614 ST---GRPSVAKTATDPVTLLADKISKNTLVVEHIDPTDCE 652
DB 486 STLDENHPSISDILDTRET-----SVNVIEGSDPEKVE 519

```

RESULT 10  
US-10-995-561-812  
Sequence 812, Application US/10995561

```
/ Publication No. US20050272054A1
/ GENERAL INFORMATION:
/ APPLICANT: CARGILL, Michele et al.
/ TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
/ TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
/ TITLE OF INVENTION: DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001559
/ CURRENT APPLICATION NUMBER: US/10/995,561
/ CURRENT FILING DATE: 2004-11-24
/ NUMBER OF SEQ ID NOS: 85702
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 812
/ LENGTH: 599
/ TYPE: PRF
/ ORGANISM: Homo sapiens
US-10-995-561-812

Query Match      3.2%; Score 111; DB 6; Length 599;
Best Local Similarity 18.6%; Pred. No. 0.46;
Matches 86; Conservative 67; Mismatches 175; Indels 134; Gaps 19;

QY 273 GSDSTSLDILKISIRGKVSILRPP--LTDEPDLDTMWOIFDIMSDDPATMGCVPN 330
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 142 GFS-SPARSLVPSSEI-LHLSLKSKVITNGPEKSSSPSSVYAAAGPRKL----- 193
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 331 TLRGAGVWPGPDVTDNFIQRHRLSYIRSHCKPNQ-----HEFMHNDKLI----- 376
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 194 ---SSGALYGRPRRTGFPSHHPOLVYISK--APSGDLIDKISELPSNKQLPFTPTTKT 247
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 377 ---TFPSANVYAIGSNKGAIVIRLNQMLPHFVOYISAASQTKRLSPKQRMGIVSSALK 433
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 248 EAKSEFLSQRYRYTPAKRKDFD-----QRIEATQTE-LSPKSELGTATKQMT 296
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 434 ELAVMRDR-----DELEDFRKY----- 453
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 297 DSENMKIKQASNCVTDAKEKIALPLBGHDSITWDEIKDLOHSSPRAMCOYSLKPPSTR 356
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 454 ---DPKDSGYISISHMKCVMENVTGLP-----PMRLRLDLAPGDSOKVNYNRITDL 504
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 357 KIYDSBELLYISFLR--DYVDEILKGLFSNRFLERLIFERIKKONKHLSEKMHLLHV 414
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 505 LDTVILELEADQMSVMDALYANKASLVAFNIIDANSGETLDEFEFTALDILVAHMPG 564
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 415 LKVDLGGCTSEENSVMQND-----VDMLVNVPDEFKAGNSBPMLKNSBVTIQQESQ 465
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 555 AYSKAEMLKCKMMDLNGGKVDLNE-----FLAARLSDLR--KEQDENIKRR 613
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 466 QYOKA-----LDMLLSAPDSENBIFPSPTEFPMPIYKSGHSGVLIQQVNDSTNLET 517
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 614 ST---GRPSVAKTATDPVTLLADKISKNTLVVEHIDIDPTDC 652
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 518 STLDENHPSISDLSLDRET-----SVNVIRGSDPEKVR 551
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 11
US-11-074-176-158
/ Sequence 158, Application US/11074176
/ Publication No. US20050250135A1
/ GENERAL INFORMATION:
/ APPLICANT: Kluenhammer, Todd R.
/ APPLICANT: Russell, William M.
/ APPLICANT: Altermann, Eric
/ APPLICANT: McAlliff, Olivia
/ APPLICANT: Perill, Andrea Azcarate
/ TITLE OF INVENTION: Nucleic Acid Sequences Encoding
/ TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore
/ FILE REFERENCE: 5051-694
/ CURRENT APPLICATION NUMBER: US/11/074,176
/ CURRENT FILING DATE: 2005-03-07
/ PRIOR APPLICATION NUMBER: 60/551,161
/ NUMBER OF SEQ ID NOS: 381
/ SOFTWARE: FastSeq for Windows Version 4.0
```

```
/ SEQ ID NO 158
/ LENGTH: 709
/ TYPE: PRF
/ ORGANISM: Lactobacillus acidophilus
US-11-074-176-158

Query Match      3.1%; Score 108; DB 7; Length 709;
Best Local Similarity 17.4%; Pred. No. 1; Mismatches 176; Indels 264; Gaps 31;
Matches 113; Conservative 97;

QY 83 DLVESFGDINVAKIELPI-RKNHIDLLIDVF-RKKRGRRLHPKYVALILREAAKSLKQLP 140
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 68 DLVER---AKNNKFPDVGDEQIDNVIEILSRKKNPV----- 104
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 141 NISPVSTAVSOQV-----TYCGDLHGGLDUL-LVHLKNGLPSNSNRYVFGDVFDRGK 193
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 105 LIGPAGVGTISIVGGLAERIAAGNVPACVAMNHIIISVINIMVAGSS---LRGSEERLK 161
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 194 RGLBVLTLILSLYLAFPNAVFLNRGNHDSVNNARVGFIRVEYSKYPNNHKKIILAFIDV 253
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 162 KYID----- 180
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 254 YRMPLGSLVLSNRYLIVHGSPSDSTSLDILKISIRGKVSILRPLTDGE-----PLDKT 308
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 181 HN-----LVGAGSTDSN-----NNGDAANILKPLAAGELKLGATTTTS 220
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 309 EMQIFDIMSDDP-----QATMGCVPTTLGAGVWPGPDVTDNFIQRHRLSYIRSHCK 362
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 221 EFORI---EKDPALSRBPQAVVPEPST-----DVAIKLBIKKKYEDVYHHV- 265
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 363 KPNGHFHFHNDKITTIFASNYVYAIGSNKGAIVIRLNQMLPHFVOYISAASQTKRL----- 418
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 266 ---KYTDDSLKLAVERISERYI---QGRY-----LPDKAIDMLBAGAKKALLVOP 309
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 419 ---SFKQRMGIVSSALK-----ELAVMRDRHDELBDFRKYDPKDSGYISISHMC 467
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 310 TDBKSLKNQISALBAKKAABEDYDKAAIEKKIATLSGQKSVDSKNTPEVTAKOY 369
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 468 KYMENVTGLSPMRLRLKRLAPGDSOK-----VNYRITDLDTVILEAB- 514
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 370 QIIBOKTIPMS-BLHABQAKNLDLAKKLKKNVLDODRAIDVI-TDAIARQIIPKSDR 427
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 515 -----ADGMSVMDALYANKASLVAFNIIDANSGETLDEP-ETV 554
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 428 PTGSEFLTGPTGVGKTELAKQALAI--LFGNKEHLI-----RLDMSEIQDEMA 473
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 555 IDLVAHMPG-----AYS-----KA-----EMLEKCRMDL 580
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 474 VNKLIGSAPGVYVGGGQLTEKVRHQPYSLILFDEIEKAMPQVENALLQIMDDRLTDA 533
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 581 NG-----DKVDLNEFLA-----FRLSDLR 602
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 534 QGRVSPFDITILIMTSNAGFSDKLEBQKVDKDLISALBNYFRPEFLNR 583
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 12
US-11-202-516-10
/ Sequence 10, Application US/11202516
/ Publication No. US2006008465A1
/ GENERAL INFORMATION:
/ APPLICANT: Steinna, Lucilla
/ APPLICANT: Mouritsen, Soren
/ APPLICANT: Gautam, Anand
/ APPLICANT: Dalum, Iben
/ APPLICANT: Haaning, Jesper
/ APPLICANT: Leach, Dana
/ APPLICANT: Nielsen, Klaus
/ APPLICANT: Karlsson, Gunilla
/ APPLICANT: Raamsen, Peter
/ TITLE OF INVENTION: Novel Methods for Therapeutic Vaccination
/ FILE REFERENCE: 4614-0107PUS2
/ CURRENT APPLICATION NUMBER: US/11/202,516
/ CURRENT FILING DATE: 2005-08-11
```

```
; PRIOR APPLICATION NUMBER: US 09/806,703
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: PCT/DK99/00525
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: DK 1998 01261
; PRIOR FILING DATE: 1998-10-05
; PRIOR APPLICATION NUMBER: US 60/105,011
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 694
; TYPE: PRF
; ORGANISM: Mus musculus
US-11-202-516-10

Query Match      3.1%; Score 107.5; DB 7; Length 694;
Best Local Similarity 19.9%; Pred. No. 1.1;
Matches 109; Conservative 72; Mismatches 194; Indels 173; Gaps 29;

QY 38 EVASEDDQAELEYKFENDLIKHPQAGRKQYOGSAHVSLDDKDLVEEFG-DIVNAKI 96
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 4 EPLHELKANKIKKFLNPFTR-TPHLAGTQNNBELAQI-----HQQWKEFGDLV- 52
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 97 ELPFRKNHIDLIDVFRKKRGNRLHPKYVALLIREAAK---SLKQLP-----NIS--P 144
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 53 EL-----SHYDVLISY-----PNKTHPNYISIIINEDGNEIFKTSISEQPPPGYENISDVVP 103
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 145 VSTAVSQOYVCGDL-----HGKLDLILVHLKNGLPSSENPVYFNGDFVDRGRGLEGVL 200
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 104 PYSAFSPQSTPEBDLVYVNTARTEDFPKLEREMKISCSGKIYARVYKGRGN----- 156
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 201 LLLSLYAFNNAVFLNRGNHEDSVNNARYGFIREVESKYPRNHKRIIAFIDEVYRW-LPL 259
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 157 MYKNAQLAGAKGMIL-----YSDP---ADY-FVPAVKS-YPDG-----WNLPG 194
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 260 GSYLNSRVLLVHGSGFSDTSLDLIKSIDRGKYVSIIRPPLTDSGEPDLKTEWQO----- 312
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 195 GGVQRGNVNLNGA-----GDLPTGYPANBEHAYRHELTNAV 232
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 313 -----IFDIMSPDQATMGCVNPLTRGAGV-----FGPDVTDNFLQHRRLSYVIR 358
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 233 LPSIPVHPIGYDDAQKLEHMGAPAPDSSMKGKLPVNVGPGFAGNF-----S 252
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 359 SHECKENGHEFMDNKITITIFASNYVYAIISNGKAYIRLNNQIMPHFVOYISAASQTKRL 418
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 283 TOKVKNHHSYTKVTRI-----YNVIGTLKGA-----LEPD--RYVILGHRDAM 325
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 419 SF-----KQRMGIYESSALKELAVRMRDHDELEDFRKYDPPDSGYISISHCKYMEV 473
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 326 VFGIDPQSGAAVVEIYVRSFGLTKKGRPRRTIIFASWDABEFGLGSTEWABEHS-- 383
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 474 TKLGLPWRLLR-----DKLAPGDSQKN-----YNRLLDLDDTVILLEAA 515
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 384 -----RLQERGVAYINADSSIEGNYTLRVDCPLMTSLVYNTLKEIQQSPDEGFE--- 433
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 516 DGMVMDA 523
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 434 -GKSLYDS 440
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
US-11-202-516-8
; Sequence 8, Application US/11202516
; Publication No. US20060008465A1
; GENERAL INFORMATION:
; APPLICANT: Steinaa, Lucilla
; APPLICANT: Mouritsen, Soren
; APPLICANT: Gautam, Anand
; APPLICANT: Dattam, Iden
; APPLICANT: Haaning, Jepper
; APPLICANT: Leach, Dana
; APPLICANT: Nielsen, Klaus
```

```
; APPLICANT: Karlsson, Gunilla
; APPLICANT: Rasmussen, Peter
; TITLE OF INVENTION: Novel Methode for Therapeutic Vaccination
; FILE REFERENCE: 4614-0107PUS2
; CURRENT APPLICATION NUMBER: US/11/202,516
; CURRENT FILING DATE: 2005-08-11
; PRIOR APPLICATION NUMBER: US 09/806,703
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: PCT/DK99/00525
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: DK 1998 01261
; PRIOR FILING DATE: 1998-10-05
; PRIOR APPLICATION NUMBER: US 60/105,011
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 752
; TYPE: PRF
; ORGANISM: Mus musculus
US-11-202-516-8

Query Match      3.1%; Score 107.5; DB 7; Length 752;
Best Local Similarity 19.9%; Pred. No. 1.2;
Matches 109; Conservative 72; Mismatches 194; Indels 173; Gaps 29;

QY 38 EVASEDDQAELEYKFENDLIKHPQAGRKQYOGSAHVSLDDKDLVEEFG-DIVNAKI 96
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 62 EPLHELKANKIKKFLNPFTR-TPHLAGTQNNBELAQI-----HQQWKEFGDLV- 110
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 97 ELPFRKNHIDLIDVFRKKRGNRLHPKYVALLIREAAK---SLKQLP-----NIS--P 144
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 111 EL-----SHYDVLISY-----PNKTHPNYISIIINEDGNEIFKTSISEQPPPGYENISDVVP 161
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 145 VSTAVSQOYVCGDL-----HGKLDLILVHLKNGLPSSENPVYFNGDFVDRGRGLEGVL 200
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 162 PYSAFSPQSTPEBDLVYVNTARTEDFPKLEREMKISCSGKIYARVYKGRGN----- 214
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 201 LLLSLYAFNNAVFLNRGNHEDSVNNARYGFIREVESKYPRNHKRIIAFIDEVYRW-LPL 259
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 215 MYKNAQLAGAKGMIL-----YSDP---ADY-FVPAVKS-YPDG-----WNLPG 252
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 260 GSYLNSRVLLVHGSGFSDTSLDLIKSIDRGKYVSIIRPPLTDSGEPDLKTEWQO----- 312
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 253 GGVQRGNVNLNGA-----GDLPTGYPANBEHAYRHELTNAV 290
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 313 -----IFDIMSPDQATMGCVNPLTRGAGV-----FGPDVTDNFLQHRRLSYVIR 358
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 291 LPSIPVHPIGYDDAQKLEHMGAPAPDSSMKGKLPVNVGPGFAGNF-----S 340
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 359 SHECKENGHEFMDNKITITIFASNYVYAIISNGKAYIRLNNQIMPHFVOYISAASQTKRL 418
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 341 TOKVKNHHSYTKVTRI-----YNVIGTLKGA-----LEPD--RYVILGHRDAM 383
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 419 SF-----KQRMGIYESSALKELAVRMRDHDELEDFRKYDPPDSGYISISHCKYMEV 473
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 384 VFGIDPQSGAAVVEIYVRSFGLTKKGRPRRTIIFASWDABEFGLGSTEWABEHS-- 441
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 474 TKLGLPWRLLR-----DKLAPGDSQKN-----YNRLLDLDDTVILLEAA 515
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 442 -----RLQERGVAYINADSSIEGNYTLRVDCPLMTSLVYNTLKEIQQSPDEGFE--- 491
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 516 DGMVMDA 523
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 492 -GKSLYDS 498
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
US-10-793-626-2964
; Sequence 2964, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
```

```

/ TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
/ FILE REFERENCE: PU3480US
/ CURRENT APPLICATION NUMBER: US/10/793,626
/ CURRENT FILING DATE: 2004-03-04
/ PRIOR APPLICATION NUMBER: 60/164,258
/ PRIOR FILING DATE: 1999-11-09
/ NUMBER OF SEQ ID NOS: 4472
/ SOFTWARE: Patent In Ver. 2.1
/ SEQ ID NO 2964
/ LENGTH: 5024
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: synthetic
/ FEATURE:
/ OTHER INFORMATION: amino acid sequence
/ NAME/KEY: MOD_RES
/ LOCATION: (5024)
/ OTHER INFORMATION: variable amino acid
US-10-793-626-2964

```

Query Match 3.1%; Score 105; DB 6; Length 5024;

Best Local Similarity 20.0%; Pred. No. 30;

Matches 149; Conservative 105; Mismatches 274; Indels 218; Gaps 36;

```

QY 53 NDLIKHM-----PQAGRKQYQGSAAHVSVDLDDKDLVDFPGDIYNAKIS 97
DB 3605 NDEIKMNTYLNAAQKOSIDKMTSHALRTBYKQLDQAKTLDEAMSLSDKQVITDTT 3664
QY 98 LP-----IRKNHIDLID---VERKRGKRLHPRVYALLIREAKSLKOLPNISPV 146
DB 3665 LPVYTBASBDEKKEKQVQTSQAQALIDKINGSVNSLDQVRQALQQLQASBNLDGQRYE 3724
QY 147 TAVSQQVTCGDLHGK---LDDLVLVLAHKGSLPSSSPYFNGDFVDRGRLGLVLLILLS 204
DB 3725 EA-----KXHANQITDQ---THLSLQOQT-----AESVYNAATKLER 3760
QY 205 LVYLAFFNAVFLNRG-----NHSDVYNNARYGFIREVSKYPRHKKILAFIDEVYM 256
DB 3761 IATASNALALANKVMKLEQFTHADSIENS-----DNTRQADDDKIIAYDVLBLENG 3812
QY 257 -----LPLGSVLSRVLIIVHGF-----SDSTSLDLKSIDR--GKYVSIL 295
DB 3813 QDIQSNATONBEAKQALQOLINAEISL--NGFERLHARPRALLEYIKSLKINNAAQKSL 3870
QY 296 RPPULTGEBLDEK-----TEKQQLFDIM-----NSDPQATMGCV----- 329
DB 3871 EDKVTQSHDLLEHLVNSGTMLNDIMGELANAIYNNYAPYKASINYINADMLRKDNFTQ 3930
QY 330 -----NTLFGAGVWFGPDVT---DNFLOHRHLSYVIRSHCKPNGHEFMHDKIT- 376
DB 3931 AINNADALANKTQGMIDFRALITFYKODIFKTYDALNGIERLTAASKAKELIDSLKPIFN 3990
QY 377 -TIFASNTYVAIGSNKAYI-----RLNNQMLPHFVQ-----YISAA 412
DB 3991 KQGFTHANDEIMVNTNSIAQLSRIVNOAFPLDAMKSLRDELINNAQAP--VQASSNYTND 4048
QY 413 SQCHKLSFKQRMKIVSSSALKELAVMRDRDELLEBFKY--DPRD--SGYISISH-W 466
DB 4049 E-----DLKQOQFHAISMARAKVLAKEKNGKLDIEIQELKQVIEDTKDLNGIQRSLSKK 4103
QY 467 CKYMNVTYKLGLPWRLRLDLKAPGTDQK---VNNRRTLDLLDTDVIIIEABADGMSV- 521
DB 4104 AKKIQVVOGLSYI-----NDQKRHTIASNTIHNSDDLSLANTLSKASDLDNAMK 4152
QY 522 ---DALYANKASIVAFNIIDADNGSEITLDEF---ETAIDLLVAHMPGAYSKAEMLERK 575
DB 4153 DLKDTLESNSTSVNSVYINADKNLQIEFDLALQOASATSSKTSNPAITIEVGLSLQA 4212
QY 576 RMMD-----LNGDQKV--DLNEPLFAFR--LSDLRKQOQDENIRRSSTGSPVAKTYT--- 625
DB 4213 -YDTYONALNGEORLATEKSKDKLTKGLDLNKAQLBD-----VTNKVNSAMTILTEIS 4265

```

```

QY 626 -----DPTVLADKISKNTLV 641
DB 4266 QLTOSTLKDNDKMLLRDKL--KTLV 4289

```

RESULT 15

US/11/062

Sequence 4, Application US/11062471A

Publication No. US20050255093A1

GENERAL INFORMATION:

APPLICANT: SHONE, Clifford Charles

APPLICANT: SUTTON, John Mark

APPLICANT: HALLIS, Bassem

APPLICANT: SILMAN, Nigel

TITLE OF INVENTION: Delivery of Superoxide Dismutase to Neuronal Cells

FILE REFERENCE: 1581.0800001

CURRENT APPLICATION NUMBER: US/11/062,471A

CURRENT FILING DATE: 2005-02-22

PRIOR APPLICATION NUMBER: 09/831,050

PRIOR FILING DATE: 1999-11-05

PRIOR APPLICATION NUMBER: PCT/GB99/03699

PRIOR FILING DATE: 1999-11-05

PRIOR APPLICATION NUMBER: GB 9824282.9

PRIOR FILING DATE: 1998-11-05

NUMBER OF SEQ ID NOS: 11

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 4

LENGTH: 1070

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Construct comprising Mn-SOD from B. stearothermophilus, a linker,

US/11/062,471A-4

Query Match 3.0%; Score 104.5; DB 7; Length 1070;

Best Local Similarity 18.7%; Pred. No. 3.4;

Matches 116; Conservative 98; Mismatches 182; Indels 223; Gaps 33;

```

QY 10 IFIOKRY-----RRQARERMRRCMQIPLNLS-----YASQOQABLYKRFND-- 54
DB 478 LIVAQMISTVNTQFTTIEGMYKALNYQA-QALBRIIKRYNYISEKKSNTINDFNIN 536
QY 55 --LIGMPQAGRKQYQGSAAHVSVDLDEK---DVLVEFGDIYNAKIELPIRKNHIDLID 110
DB 537 SKLNGCIGQALDININFPNGSVSYLMKMKIPLAVKLTDPNT-----LKNGLANYIDE 591
QY 111 VERKKRGRLHPKRYVALILREAKSLKOLPNISPVSTAVSQQVTCGDLHGKL----- 163
DB 592 -----NKLX--LIGSAVEYKSKVNYKLTIMPDLSIYTDITLIEFNYKNSGILLNN 642
QY 164 -----DOLLVLAHNG-----LPSGSPN-----YFNGDF 188
DB 643 IILNLRVYDNMLIDISGYGAKVEYVYGVELNDKQFKLTSSANSKIRVTQONQIIFNSVF 702
QY 189 VDRGRGLSEVLLLLSLVLAFP---NAVFLNRGNHSDVMNA----- 227
DB 703 LD-----FSTVSPMIRIPKYQNDGIQNYIHIEYITIIINCKNNNSGKISIRGNRIIW 752
QY 228 -----RYGFIREVSKYPRHKKILAFIDEVYM--LPLGSVLSRVLIIVH 271
DB 753 TLIDINGKTKSVFFRYN--IREDISY-----INRWFTYITNNLNNAKIYIN 798
QY 272 GGFSPSTSL-DLKSIDNGKYVSIIRPPLTDEBPLDKTE--WQQLFDIMSDPQATMGCV 328
DB 799 GRLBSNTIKIDIRVIANGELIFKL-----DQD-IDRQFIPMKYFSI----- 840
QY 329 PTLTLAGVWFGPDVTDNFLOHRHLSYVIRSHCKPNGHEFMHND-----NKLITIFS 380
DB 841 -----FTTLESQSNIBER---YKIQSIS-----EYAKDWGMPPLMINKKTYMKN 881
QY 381 ASNYVAIGSNKAYIRLNNQMLPHFVOYISAASQTKRLSFQKRMKIVSSSALKELAVMR 440
DB 882 A-----GNKSYIKLKDKSPVGEILTRFSKYNQNSKYINRYDLYIGR-----KFLIRRK 929

```



```

Qy 441 HHBRLAEDEF-RKXDPKDSGISISHWCKWMENTYKGLPRLRLDK-----LAP 489
Db 930 SNSGSINDIYRKED-----YIYDFF-----NLQSRKVYTYTKYFKCKBEKGLFIAP 976
Qy 490 GTDSQKVNYNRTLDDLTDT 508
Db 977 ISDSDEF-YN-TIQIKED 993

RESULT 16
US/11/062
; Sequence 7, Application US/11062471A
; Publication No. US20050255093A1
; GENERAL INFORMATION:
; APPLICANT: SHONE, Clifford Charles
; APPLICANT: SUTTON, John Mark
; APPLICANT: HALLIS, Basam
; APPLICANT: SILMAN, Nigel
; TITLE OF INVENTION: Delivery of Superoxide Dismutase to Neuronal Cells
; FILE REFERENCE: 1501.0800001
; CURRENT APPLICATION NUMBER: US/11/062,471A
; CURRENT FILING DATE: 2005-02-22
; PRIOR APPLICATION NUMBER: 09/831,050
; PRIOR FILING DATE: 1999-11-05
; PRIOR APPLICATION NUMBER: PCT/GB99/03699
; PRIOR FILING DATE: 1999-11-05
; PRIOR APPLICATION NUMBER: GB 9824282.9
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 1095
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Construct comprising a mitochondrial leader sequence from Human P
US/11/062,471A-7

Query Match 3.0%; Score 104.5; DB 7; Length 1095;
Blast Local Similarity 18.7%; Pred. No. 3.6;
Matches 116; Conservative 98; Mismatches 182; Indels 223; Gaps 33;

Qy 10 IFIQKWY-----RRHQARREWRRCNMQIFQNLB-----YASEODAEIYKFPND- 54
Db 503 LIVQWMLSTVNTQFTTKEGMYKALNYQA-QALEETITRRNYITSEKKSININDFENIN 561
Qy 55 --LTKHMPQAGRKNOYQSAHVSVLDDK--DDVEERFGDVIYNAKIEPIRKNHIDLID 110
Db 562 SKLNGINGQALDINNPFINGCSVSLMKKKMPLAVEKLLDPNT-----LKNLLNLYIDE 616
Qy 111 VPRKKGRRLHPKYVALILRLPAKSLKQLPNISPVSTIVSQQVYCGDLHGKL----- 163
Db 617 -----NKLV--LIGSAIEYKSKYNYKIKTIMPDLSIYNTDTLLIEMFKNYSEILNN 667
Qy 164 -----DDLVLVHLKNG-----LPSSNP-----YVFNQDF 188
Db 668 IILNRLRYKDNLLIDLSGYGAKYEVYDGYELADKXQPKLITSSANSKIRVTONONITFNSVF 727
Qy 189 VDRGRGLEVLILLLSLYLAP--NAVELNRGNHEDSVANA----- 227
Db 728 LD-----FSVSPWIRIPKYKNDGIQVYIHNEYIILNCKNNKSGMKLSIRGNRIIW 777
Qy 228 -----RYGRIFVESKYPRNHRIILAFIDEVYRW--LPLGSVLNLSVLIVH 271
Db 778 TLIDINGTKSVFPEYN-IREDISY-----INRWFPVITITNNLNNAKIYIN 823
Qy 272 GGFSDSTSL-DLIRKSIDRGKYVSLIRPPLDGEPLDKTE--WQQLFDITWSDPQATMGCV 328
Db 824 GGLBSSTLIDKQIRVIANGELIFKL-----DQD-IDRQPIWMKXFSL----- 865
Qy 329 PNTLRGAGVWGGPDVDTNDFLQHRHLSTYIRSHCEKPNCHHEFMD-----NKLTITFS 380

```

```

Db      866 -----FNTLSGSNIIEER---YKIGSYS-----EYLXDFWGNLPLNNKBEYMFN 906
Qy      381 ASNTYAIGSNKGAVIRLNNOLMHPFVOYISAASTKELSPKQRMGIYESSALKELAVRM 440
Db      907 A-----GKNKSYYIKLKQSPVGEILTRSKYCNQNSKYINRDLYIGB-----KFILRRK 954
Qy      441 DHRLEADEF-RKYDPDPSGYISISHMCKWMENTYKLGLPWRLLRDK-----LAP 489
Db      955 SNSGSINDDIYRKD-----YIYDFP-----NINQERWYTYTYFKKBEKFLFAP 1001
Qy      490 GTDSQKVNRYRTLDLPTD 508
Db      1002 ISDSEDF-YN-TIQIKEDY 1018

RESULT 17
US-10-118-590-18
; Sequence 18, Application US/10118590
; Publication No. US20050277761A1
; GENERAL INFORMATION:
; APPLICANT: KENNETH RHODES, MARIA BETTY, HUAI-PING LING, AND FRANK AN
; TITLE OF INVENTION: POTASSIUM CHANNEL INTERACTORS AND USES THEREFOR
; FILE REFERENCE: NMI-070
; CURRENT APPLICATION NUMBER: US/10/118, 590
; CURRENT FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 2002-04-08
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 270
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-118-590-18

Query Match      3.0%; Score 103; DB 6; Length 270;
Best Local Similarity 25.7%; Pred. No. 0.59;
Matches 59; Conservative 25; Mismatches 58; Indels 88; Gaps 13;

Qy      439 MRDR-----DELEDFRKYPDPSGYISISHMCKWMENTYKLGLPWRLLRDKLA----- 488
Db      64 LRPRRPRLPDPSVEDF-----ELSTVCHRPBGLEJOLQBOTKPTRRRLQYLYR 112
Qy      489 -----PGTDSQKVNRYRTL-DLPTD-----VILBAADGMS 519
Db      113 GPKNRCSPGIWENPFGQIYSQFPFGDSS--NYATFLPNAFDINHDGSVSEDFVAGLS 170
Qy      520 VM-----DALYANKASLVALENIIDANSGETLDEFETAILD--VAHMPGAYSXAEW 571
Db      171 VILRGITDRL--NMA-----FNLJYDNKDCGCIKBEW--LDIMKSIYDMMGKTYPAL 220
Qy      572 LEKC-----RMNDLNGDKYDLNLEPFLSLDLHREKQODENIR 612
Db      221 REBAPREHVESFFQKMDRNKQGVITIEFIESC-----QQDENIMR 261

RESULT 18
US-10-055-877-140
; Sequence 140, Application US/10055877
; Publication No. US20050288241A1
; GENERAL INFORMATION:
; APPLICANT: Decristofaro, Marc
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Miller, Charles
; APPLICANT: Tchermey, Velizar
; APPLICANT: Zhong, Wei
; APPLICANT: Anderson, David
; APPLICANT: Ballinger, Robert
; APPLICANT: Gerlach, Valerie
; APPLICANT: Spytek, Kimberly
; APPLICANT: Ratelli, Luca
; APPLICANT: Kerkuda, Ramesh
; APPLICANT: Guo, Xiaojia

```

```

APPLICANT: Zerhusen, Bryan
APPLICANT: Andrew, David
APPLICANT: Metzes, Peter
APPLICANT: Paccureajan, Meera
APPLICANT: Burgess, Catherine
APPLICANT: Eisen, Andrew
APPLICANT: Wolenc, Adam
APPLICANT: Baumgartner, Jason
APPLICANT: Shinkets, Richard
APPLICANT: Gusev, Vladimir
APPLICANT: Vermet, Corine
APPLICANT: Taupier Jr., Raymond
APPLICANT: Pena, Carol
APPLICANT: Shenoy, Suresh
APPLICANT: Li, Li
APPLICANT: Caeman, Stacie
APPLICANT: Boldog, Ference
TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
FILE REFERENCE: 21402-251
CURRENT APPLICATION NUMBER: US/10/055,877
CURRENT FILING DATE: 2002-01-22
PRIOR APPLICATION NUMBER: 60/262,892
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: 60/263,598
PRIOR FILING DATE: 2001-01-23
PRIOR APPLICATION NUMBER: 60/263,799
PRIOR FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: 60/264,117
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/264,139
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/264,478
PRIOR FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 60/263,351
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/272,870
PRIOR FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: 60/275,990
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/275,927
PRIOR FILING DATE: 2001-03-14
Remaining Prior Application data removed - See file Wrapper or PALM.
NUMBER OF SEQ ID NOS: 512
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 140
LENGTH: 1765
TYPE: PRT
ORGANISM: Mus musculus
US-10-055-877-140

Query Match      3.0% Score 103; DB 6; Length 1765;
Best Local Similarity 21.1%; Pred. No. 9.4;
Matches 107; Conservative 54; Mismatches 177; Indels 168; Gaps 26

QY 37 LEVASEDQDAELKRFNDLIKMPDAAGKQYQGSAAVSVLDDQDLVEER---GDIVN 93
DB 453 LHMARASQGAVERV---LVDDGAQVBAKADDQTPPLHISARLGRADIVQQLDQGSAPN 509
QY 94 A-----KIEPLIKNHIDLLIDIVFRKKGRNRLHPKVVALIRE--AKSLKOLPNTSP 144
DB 510 AATTSIGYPLHLARREGHD-----VAAFLDQHASLSITTYKKGFTP 551
QY 145 VSTAVSQGVTCGDLAGKLDLVLVLAHNGVLPSASNPFVFNQDFVDRGKGL----- 196
DB 552 LHVAAK-----YGLLEVASLLLOKASAPDA-----QKSGITPLHVAH 590
QY 197 -----EVLILLISLYIAFNNAVFLN-----KGNHDSVNN-ARYGF-----IR 233
DB 591 YDNQVALLLLLD-QGASPHAAKNGYTPPLHIAKKQCDIATSLLEYADANAVTRQGIA 649
QY 234 EYBSGYPNHNRKRIILAFIDEVYRWLPLGSVLSNRVLVH-CGFSDSSTSLDKSIDRGKV 292
DB 650 SVHLAAQGRH-----VDMV-----SLLSRNANVLSNKGSLTPLHLAAQEDRVAVA 696

```

```

Oy      293  SILLRPPLTQSEPLDKTKEWQIQIPIDIMSDQATWGCVPNTLRGAGWPFQGDVTDNPLQGBR 352
Db      697  EVL-----VNGQAHY-----DAQTKMGITP---LHVQCHYGIKIVNPLLOHS 736
Oy      353  LSYVIR8-----HRCQNGHEFMHNDKITITIPASNYTAIGSNKGAyi--RLNNQOLMP 403
Db      737  AKVNAKTKNGYTAHQAAQGGHTHI--INVLLQNNASPNELTVNGNTALAIABRLG----- 790
Oy      404  HPOYIYSAA8-----QTRGLSPKQRMGIVESALKELAYRRRHRHRLDEBFKX 453
Db      791  ----YISVDTTLKVYTEEIMWTITITITEKHKMNVPET-----NNEVLDMSDDEVKRA 837
Oy      454  DP-----KDSGYISISH-----WCKV 469
Db      838  SAPEKSLDSEYISIDGEBGDKCTWFKI 863

RESULT 19
US-10-821-234-1535
/ Sequence 1535, Application US/10821234
/ Publication No. US20050255114A1
/ GENERAL INFORMATION:
/ APPLICANT: Ladac, Ivan
/ APPLICANT: Scache-Crain, Birgit
/ APPLICANT: Andermant, Susan
/ APPLICANT: Tang, Y. Tom
/ TITLE OR INVENTION: Methods for diagnosis and treatment of Preeclampsia
/ FILE REFERENCE: 821A
/ CURRENT APPLICATION NUMBER: US/10/821,234
/ CURRENT FILING DATE: 2004-04-07
/ PRIOR APPLICATION NUMBER: US 60/462,047
/ PRIOR FILING DATE: 2003-04-07
/ NUMBER OF SEQ. ID NOS: 1704
/ SOFTWARE: PC_SEQ_genes Version 1.0
/ SEQ ID NO 1535
/ LENGTH: 149
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-821-234-1535

Query Match 3.0%; Score 102; DB 6; Length 149;
Best Local Similarly 28.1%; Pred. No. 0.3;
Matches 39; Conservative 20; Mismatches 58; Indels 22; Gaps 6;

Oy      483  LRDLKAGTDSOKVNNRTLDLDTDVILLEABDGM-----SVMDALYAN--KASLYA 533
Db      1 MADOL--TEBQIAFFKAFSLPKD-----GDGTYTTKEIGTVKRSIQONPTRELEDD 51
Oy      534  IFNIIDADNSGEITLDEFETAIIDLVAHMPGAYSKAEMLEKGRMDLNDGKVDLNEFLE 593
Db      52  MINEVVDADNGTIDTDPFELT---MMARKKDDTSEEBEINEAPFVFXDONGYISAAELNH 108
Oy      594  AF-RISDLHRKEQODENIR 611
Db      109  VMTNLGKRLTDEEVDEWIR 127

RESULT 20
US-11-124-368A-268
/ Sequence 268, Application US/11124368A
/ Publication No. US20050287559A1
/ GENERAL INFORMATION:
/ APPLICANT: Michele Cargill
/ APPLICANT: James J. Devlin
/ APPLICANT: May Luke
/ TITLE OR INVENTION: Genetic Polymorphisms Associated with
/ TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
/ FILE REFERENCE: CU001524
/ CURRENT APPLICATION NUMBER: US/11/124,368A
/ CURRENT FILING DATE: 2005-05-09
/ PRIOR APPLICATION NUMBER: US 60/568,845
/ PRIOR FILING DATE: 2004-05-07

```

```

      PRIOR APPLICATION NUMBER: US 60/625,996
      PRIOR FILING DATE: 2004-11-09
      NUMBER OF SEQ ID NOS: 21112
      SOFTWARE: FastSeq for Windows Version 4.0
      SEQ ID NO 268
      LENGTH: 149
      TYPE: PRT
      ORGANISM: Homo sapiens
      US-11-124-368A-268

Query Match
Best Local Similarity 3.0%; Score 102; DB 7; Length 149;
Matches 39; Conservative 20; Mismatches 58; Indels 22; Gaps 6;

Cy 483 LRDKLAPGTDTSQKYNVNRITLDLDTDVILAEADGM-----SVNDALYAN--KASLVA 533
Db 1 MAOQL---TEBQIAEFKAPSLFDKD-----GDGTTTKEIGTAVRSIGQNPTEAEIGD 51

Cy 534 IFNIIIDANSGETITLDEFETAIIDLVAHMPGAYSKAEMLEKCKMMDLNGDKVDLNEPFE 593
Db 52 MINEVVDADGNGTIDPEPFLT---MMARKMDTSEBEIRAEAFVFDKDGNGYISAABLH 108

Cy 594 AF--RLSDLRKEQDENIR 611
Db 109 VMTNLGSKLTDEEVDENIR 127

RESULT 21
US-11-124-368A-269
Sequence 269, Application US/11124368A
Publication No. US20050287559A1
GENERAL INFORMATION:
APPLICANT: Michele Cargill
APPLICANT: James J. Devlin
APPLICANT: May Luke
TITLE OF INVENTION: Genetic Polymorphisms Associated with
TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
FILE REFERENCE: C1001524
CURRENT APPLICATION NUMBER: US/11/124,368A
CURRENT FILING DATE: 2005-05-09
PRIOR APPLICATION NUMBER: US 60/568,845
PRIOR FILING DATE: 2004-05-07
PRIOR APPLICATION NUMBER: US 60/625,936
PRIOR FILING DATE: 2004-11-09
NUMBER OF SEQ ID NOS: 21112
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 269
LENGTH: 149
TYPE: PRT
ORGANISM: Homo sapiens
US-11-124-368A-269

Query Match
Best Local Similarity 3.0%; Score 102; DB 7; Length 149;
Matches 39; Conservative 20; Mismatches 58; Indels 22; Gaps 6;

Cy 483 LRDKLAPGTDTSQKYNVNRITLDLDTDVILAEADGM-----SVNDALYAN--KASLVA 533
Db 1 MAOQL---TEBQIAEFKAPSLFDKD-----GDGTTTKEIGTAVRSIGQNPTEAEIGD 51

Cy 534 IFNIIIDANSGETITLDEFETAIIDLVAHMPGAYSKAEMLEKCKMMDLNGDKVDLNEPFE 593
Db 52 MINEVVDADGNGTIDPEPFLT---MMARKMDTSEBEIRAEAFVFDKDGNGYISAABLH 108

Cy 594 AF--RLSDLRKEQDENIR 611
Db 109 VMTNLGSKLTDEEVDENIR 127

RESULT 22
US-10-821-234-935
Sequence 935, Application US/10821234
Publication No. US20050255114A1

```

```
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmant, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pc_seq_genes Version 1.0
; SEQ ID NO 935
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-935

Query Match      3.0%; Score 102; DB 6; Length 171;
Beat Local Similarity 28.1%; Pred. No. 0.36;
Matches 39; Conservative 20; Mismatches 58; Indels 22; Gaps 6;

QY    483 LRDLAPETDSOKVNNRITLDDTLVDVILEADGM-----SYMDALYAN--KASLYAIRIIDD 533
       :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~
DB     23 MADQL---TEBQIAEFKEAFSLFDKD-----GDTITTKEIGTVRSISGONPTAEILQD 73
OY     534 IFNIIDANSSEITLDEPETAIDLVAHMPGAYSKAEMLEKCRMDLNDGKVDFLNFELE 593
Db      74 MINVEDADNGNTTTPPEFLT---MARRKKOTDSEERIEARVRVDKGNGIISAELRH 130
OY     594 AF-RLSDLRRKQODENIR 611
       |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      131 VMTNLGEKLTDREVDENIR 149

RESULT 23
US-11-065-943-46
; Sequence 46, Application US/11065943
; Publication No. US20050250131A1
; GENERAL INFORMATION:
; APPLICANT: JESTIN, JEAN-LUC
; APPLICANT: VICHER-GUERRE, SOPHIE
; APPLICANT: PEREIRA, STEPHANE
; TITLE OF INVENTION: METHODS FOR OBTAINING THERMOSTABLE ENZYMES, DNA POLYMERASE I
; TITLE OF INVENTION: VARIANTS FROM THERMUS AQUATICUS HAVING NEW CATALYTIC ACTIVITIES
; TITLE OF INVENTION: METHODS FOR OBTAINING THE SAME, AND APPLICATIONS OF THE SAME
; FILE REFERENCE: 266426USOXCP
; CURRENT APPLICATION NUMBER: US/11/065,943
; CURRENT FILING DATE: 2005-02-25
; PRIOR APPLICATION NUMBER: US 10/787,219
; PRIOR FILING DATE: 2004-02-27
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 46
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-065-943-46

Query Match      2.9%; Score 101.5; DB 7; Length 148;
Beat Local Similarity 28.2%; Pred. No. 0.32;
Matches 37; Conservative 18; Mismatches 57; Indels 19; Gaps 5;

QY     491 TDOSKVVNRIPLDLLDPVILEADGM-----SYMDALYAN--KASLYAIRIIDD 541
       |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~
DB      5 TEBDIAREFAFSLPDKD-----GDGITTTKEIGTVRSISGONPTAEILODMINEVDAD 58
OY     542 NSGETITDEFETAIDLVAHMPGAYSKAEMLEKCRMDLNDGKVDFLNFELEAF-RLSDL 600
Db      59 GNGITIDPEFLT---MARARKOTDSEERIEARVRVDKGNGIISAELRHVNLTNGEK 115
OY     601 HRKEQODENIR 611
```



```

? APPLICANT: CHIRON SPA
? APPLICANT: FONTANA Maria Rita
? APPLICANT: PIZZA Mariagrazia
? APPLICANT: MASTIGNANI Vega
? APPLICANT: MONACI Elisabetta
? TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
? FILE REFERENCE:
? CURRENT APPLICATION NUMBER: US/10/467,657
? CURRENT FILING DATE: 2003-08-11
? PRIOR APPLICATION NUMBER: GB-0103424.8
? PRIOR FILING DATE: 2001-02-12
? NUMBER OF SEQ ID NOS: 9218
? SOFTWARE: Seqwin99, version 1.04
? SEQ ID NO 2502
? LENGTH: 276
? TYPE: PR1
? ORGANISM: Neisseria gonorrhoeae
? US-10-467-657-2502

```

Query Match 2.9%; Score 98.5; DB 7; Length 749;  
 Best Local Similarity 19.6%; Pred. No. 5.9;  
 Matches 114; Conservative 83; Mismatches 217; Indels 167; Gaps 27;

```

OY 32 QIFQULJEASBEDQO-----ELYKFF-----DLIKMPAARKNQY-----69
Db 14 QVIDTCKKYMNEQVAFVBRAYEFANKAKAQQKRAAGQYIITHPTQVAGTLANTLGDPDT 73
OY 70 --QSSAHVSLD--DKDQVBEFGD---IVNAKIELP-----IRKNHIDLIT-109
Db 74 IAAQGLHDTVEBTPVTNDELKEKEFGEDVAFIVDQVTKLKNKEYENSHQEBFAENHRKMLLA 133
OY 110 ----DVRKKRGKNTLHPKYVALILRBAKSLKQLPNISPRVSTAVSQOYTVCGDLHGKLD 164
Db 134 MAKDLRVIMVUKLADRIHNNHTLOHLR-PDKORRIASETMDIYAPLADRLGI-GTIMWELE 191
OY 165 DL-----LVLHKNGLPSSGPNPYVFNQDVFVDROKRGLEVLILLSTYLAPPNAY 213
Db 192 DMSFTYLPBEAYTRYIYNLMDVQSRQREKTI--SDTIKTTLKTIDELGIXDYI-----242
OY 214 FLNRGNHEDSVNNAARYGFIREVESKYPNNHKKLILAFIDEVYRWLPJGSGVLSNRVLI----269
Db 243 --GRPKH-----IYSIYKKQVNGHKD---FDEIY-----DELAAYVIVKNVR 279
OY 270 -----VH-----GGFSDBTSJDLKSIDRGKYVSTILRPLTDSBPLD-KTWMQOI 313
Db 280 DCYAVLGAHTEWKMPGGRFKQYILAPKNGY-QSLHTTITIGP---GGRPLEIQIRTEQM 335
OY 314 FDIWMSDPOATWGCVPNTLRGAGVWPGRPDTNFIQORHLSYVIRSHCKPNGHEFMHDN 373
Db 336 HEV-----AEGVAAHMAKYKKGPNFGVLAITSGEGLDMVREILELKDCTKAGFEMKSV 389
OY 374 KIITFSASNY-----YAIGSNTGAY--IRLNNQIMPHFVOYI 409
Db 390 K-SDIFSRVVYFTPEKGEVYELPKGSVTLDPFAVAIHTQVGSHAVGARVNNKVLV-----442
OY 410 SAASGTKSLSEFGORNGIYVSSALKEILAVNMRHRDLF-----DERKXDPKDSGYTIS 464
Db 443 -----LDYTLRNGDVLEIMTQTNAAPSRDMADVKTSSRANKLRIRYFRQDREBSIE 494
OY 465 HWCKYMEN-VTKLGLPWRLLRDKLAPGTSOKVNNYRTLDL 504
Db 495 HGEQVAVATIRDEGLVAPDKPMDKKEHLEKLEHNHYTISEL 535

RESULT 29
US-10-995-561-772
; Sequence 772, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: C1001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 772
; LENGTH: 1538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-772

```

QY	44	-----DOAEUYXFF--NDLKJHMPQ-----AAGKJNYQGSAAH-----VSV-----77
Db	581	ILNTVLTSQIBEHKVFANBNHNRDQIILEDTQNGQLKFLSOKQDVULIKMLVSVGSRME 640
QY	78	-----LDDKDXLVEEFGDVIYNAKIE--LPIKXNIIIDLIDV-----111
Db	641	KVVGORSTIERGRSLDDAKRAKQFHEBAWKCIIDWLEDAESHLDSELSINDPKIKLOLSK 700
QY	112	---FRKRGKRLHFKYVALLIREAAKSLKQ---LPINISPVSTAVSOV-----TVCG--157
Db	701	HKEQOKTIGGK--QFVYDTTI--RTGRALKKERTILLPEBTOXLDFLGEVRDKMDTVCGKSV 757
QY	158	DLHGKLDLVLVHLKNGPLSSSNVYVNGCFVDRGK-----GLE 197
Db	758	EROKKLEAL-----LFSGGFMDLQALVDMLYKVEBQALABDQVHGDDL 802
QY	198	VLLLLLSLYLAFPNV-----FLNRGHNEDSVNNAR-----XGFIREV-----235
Db	803	LYVNLMDAHKVFQKELDKRTGTGYVLKRSGRF--LIEBSRDTTWVKGQLOLSTRMVTC 861
QY	236	-----ESKYPNRHRIILAFIDVY--RNLPLGASVLNSVVLVHGGSFSDSTSLLIKSI 286
Db	862	KLSVSKOSRLEQALKQAEFRDVTYHMLLEMLSEA---BQTLRFRGALPDDTEA--LOSLI 916
QY	287	DRGKYSILNRPFLIDGEBPLDKTEMOQFEDIMSDPQATWG-----CVPTILGAGVWFG 340
Db	917	D-----THKEFPMKKVBEKRV--DV--NSAVALMBEVLAVLCHPCITTIKIMT-- 959
QY	341	PDVTDNFOGHRLSYVRSHCEKCPNGHREFMHNKIIYIFASNYAALIGNKGAVIRLNNQ 400
Db	960	-----IIRARFEVVLTKAK-----QHQRLETFALSB-----LVANAE 992
QY	401	LMHFVQYIGASAT-----KRISFKORMGIYESSALKELAVMRDHRD 444
Db	993	LLEBLAMTOWALETILLORDQSPIPONIDRVKAL-----IAHQTFMEEMTRKOPDVD 1045
QY	445	ELDEBFK--YDPRDSGYI--STSHMKWMENTYKGLP-----479
Db	1046	RVTYTYRKNIETPHAPFIEKSRGGKRSISOPTPPMPPILOSSEALONPRINOLSANWQ 1105
QY	480	-WRLIRDKLAPGTSQKVNRYNRITDLLDTVYLLEABD-----GMSVMDALYANKASLV 532
Db	1106	VMLALER-----QRKLN-----DALDRLEELKEFANPDFVWRKRYKMYMMNHKKSRYM 1154
QY	533	AIFNIIADNSGEITLDEFTALD--LIVAHMPGAYSAEMLEKCRMDLNGDKVNDLNEF 591
Db	1155	DFPFRIRMDQOGKITRQBF---IDGILASKFP--TTKLEMTAVADIDRQDGDYIDYEF 1209
QY	592	LEAFRLS--DLHKEQODENIRRRSTGSPSAVAKTA 624
Db	1210	VVALHPNKDAVRPTTDADKIEDETVRVAOCKCA 1243

```

OY      9 AIFIOGMYRRHQARREMOQR-----CNGQIF--QNLGVASRG- 43
          |||  |||  |||  |||
Db      524 ALBQGM--HVVSSKQREERSKLEBIALNLTAFQNSLOEFINWLTABQSNIAASPPSL 580

```

Query Match 2.8%; Score 98; DB 6; Length 5335;  
 Best Local Similarity 18.9%; Pred. No. 1.2e+02;  
 Matches 154; Conservative 103; Mismatches 265; Indels 292; Gaps 43;

```

Cy 9 AIFQKRYRHHQARREMR-----CWMQIF--ONLEVASQ-- 43
Db 4364 ALLEQK--HVSSKMEERKSKLEBALNATEFQNSIQEFINWTLAEOSLNIASPPSL 4420
Cy 44 -----DOALEYKPF--NDLIKMPQ-----AARKNOYOGSAH-----VSU----- 77
Db 4421 ILNTVLSQIEBHKVFAFENVAHNRDQIIEIDQTNQKLFQKQDVVLIKNLLVSQSRWE 4460
Cy 78 -----LDDKODLVEEFQDIYNAKIE--LPIKRNHIDLIDV----- 111
Db 4481 KVVQRSIERGRSLDDAKRAKQFHEAMKCLIDWLEDBSHIDSELSINPDKIKLOLSK 4540
Cy 112 ---FRKRGNRHLPKYVALILREAKSLKQ---LPNISPVSTAVSQV---TVCG--- 157
Db 4541 HKEFQKTLGSK-QPVYDTTI--RTGRALKERTLLPEDTQKLDNPLGVRDKMDTVCGKSV 4597
Cy 158 DLHGKLDLVLVHKNGLPSSSNPYVNGDFVDRGR-----GLE 197
Db 4598 BROHKEBAL-----LPSGQPMALQALVDMLYKVEPOLAEDQVHGDL 4642
Cy 198 VLLILSLYLAFPNV-----FLNRGNHDSVMNR-----YGFIREV----- 235
Db 4643 LVNMLDANHVFQKELGRGTGVQLKRSGR--LIENSRODTTWVKQLOELSTRMDTVCG 4701
Cy 236 -----ESKYPNRHKLIAFIDEVY---RWLPLGSVLSNRYLVHGFSDSTSLDKSI 286
Db 4702 KLSVSKOSRLEQALKQAEVFRDVTVMLEWLSA---BQTLRPRGALPDPTA--LQSLI 4756
Cy 287 DRGKYVSLRPPLDGPRLDKTEWQQLFDIMSDPOATMG-----CVPTLRGAGVWFG 340
Db 4757 D-----THKEFMKVEBKRV--DV---NSAVAMGEVILAVCHPDCITTIKRWI- 4799
Cy 341 PDVTNDFLQHRISLVYIRSHCEKPNHGFHNDKIITIFASNYAIGSNKGAIRLNQ 400
Db 4800 -----TIRARFEVUTMAK-----OHQRLLETALS-----LVANAE 4832
Cy 401 LMPHFVQYISAQOT-----KRLSPQRMGIVESSALKELAVMRDHRD 444
Db 4833 LLEBLAMIQWAEFTLLIQRODEPIPNIDRVKAL-----IAHQTFMEEMTKQDPVD 4885
Cy 445 ELDEBPRK--YDPKDSGYI--SISHMCKMENVTKGLP----- 479
Db 4886 RVTKTYRKNIIEPTHAPEIKSRSGRKSLSQTPPPMPTLISQSEAKNPRINQISARWQ 4945
Cy 480 -WRLLRDKLAPGDSQKVNVRTLLDPTVILEABD-----GMSVMDALYANKASLV 532
Db 4946 VMLLALR-----QRKLN-----DALDRLEELKEFANFPDVRKKKYMNMHKKSRVM 4994
Cy 533 AIFNIIDANSGETTLDEFEFAID-LVVAHMPGAYSAEMLBKCRMDLNDGKVDLNEP 591
Db 4995 DFFRIDKDDGKITRQBF---IDGILASKFP--TTKLEMTAVADIPDRDGDYIDYEF 5049
Cy 592 LEAFRLS-DLHRKEQDENIRRRSTGRPSVAKTA 624
Db 5050 VAALHPKDAVRPTTDADKIEDVTRQVACCKA 5083

RESULT 31
US-10-995-561-774
; Sequence 774, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEROF
; FILE REFERENCE: C0001559
; CURRENT APPLICATION NUMBER: US/10/995,561

```

CURRENT FILING DATE: 2004-11-24

; NUMBER OF SEQ ID NOS: 85702  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 774  
 ; LENGTH: 5406  
 ; TYPE: PR  
 ; ORGANISM: Homo sapiens  
 US-10-995-561-774

Query Match 2.8%; Score 98; DB 6; Length 5406;  
 Best Local Similarity 18.9%; Pred. No. 1.2e+02;  
 Matches 154; Conservative 103; Mismatches 265; Indels 292; Gaps 43;

```

Cy 9 AIFQKRYRHHQARREMR-----CWMQIF--ONLEVASQ-- 43
Db 4435 ALLEQK--HVSSKMEERKSKLEBALNATEFQNSIQEFINWTLAEOSLNIASPPSL 4491
Cy 44 -----DOALEYKPF--NDLIKMPQ-----AARKNOYOGSAH-----VSU----- 77
Db 4492 ILNTVLSQIEBHKVFAFENVAHNRDQIIEIDQTNQKLFQKQDVVLIKNLLVSQSRWE 4551
Cy 78 -----LDDKODLVEEFQDIYNAKIE--LPIKRNHIDLIDV----- 111
Db 4552 KVVQRSIERGRSLDDAKRAKQFHEAMKCLIDWLEDBSHIDSELSINPDKIKLOLSK 4611
Cy 112 ---FRKRGNRHLPKYVALILREAKSLKQ---LPNISPVSTAVSQV---TVCG--- 157
Db 4612 HKEFQKTLGSK-QPVYDTTI--RTGRALKERTLLPEDTQKLDNPLGVRDKMDTVCGKSV 4668
Cy 158 DLHGKLDLVLVHKNGLPSSSNPYVNGDFVDRGR-----GLE 197
Db 4669 BROHKEBAL-----LPSGQPMALQALVDMLYKVEPOLAEDQVHGDL 4713
Cy 198 VLLILSLYLAFPNV-----FLNRGNHDSVMNR-----YGFIREV----- 235
Db 4714 LVNMLDANHVFQKELGRGTGVQLKRSGR--LIENSRODTTWVKQLOELSTRMDTVCG 4772
Cy 236 -----ESKYPNRHKLIAFIDEVY---RWLPLGSVLSNRYLVHGFSDSTSLDKSI 286
Db 4773 KLSVSKOSRLEQALKQAEVFRDVTVMLEWLSA---BQTLRPRGALPDPTA--LQSLI 4827
Cy 287 DRGKYVSLRPPLDGPRLDKTEWQQLFDIMSDPOATMG-----CVPTLRGAGVWFG 340
Db 4828 D-----THKEFMKVEBKRV--DV---NSAVAMGEVILAVCHPDCITTIKRWI- 4870
Cy 341 PDVTNDFLQHRISLVYIRSHCEKPNHGFHNDKIITIFASNYAIGSNKGAIRLNQ 400
Db 4871 -----TIRARFEVUTMAK-----OHQRLLETALS-----LVANAE 4903
Cy 401 LMPHFVQYISAQOT-----KRLSPQRMGIVESSALKELAVMRDHRD 444
Db 4904 LLEBLAMIQWAEFTLLIQRODEPIPNIDRVKAL-----IAHQTFMEEMTKQDPVD 4956
Cy 445 ELDEBPRK--YDPKDSGYI--SISHMCKMENVTKGLP----- 479
Db 4957 RVTKTYRKNIIEPTHAPEIKSRSGRKSLSQTPPPMPTLISQSEAKNPRINQISARWQ 5016
Cy 480 -WRLLRDKLAPGDSQKVNVRTLLDPTVILEABD-----GMSVMDALYANKASLV 532
Db 5017 VMLLALR-----QRKLN-----DALDRLEELKEFANFPDVRKKKYMNMHKKSRVM 5065
Cy 533 AIFNIIDANSGETTLDEFEFAID-LVVAHMPGAYSAEMLBKCRMDLNDGKVDLNEP 591
Db 5066 DFFRIDKDDGKITRQBF---IDGILASKFP--TTKLEMTAVADIPDRDGDYIDYEF 5120
Cy 592 LEAFRLS-DLHRKEQDENIRRRSTGRPSVAKTA 624
Db 5121 VAALHPKDAVRPTTDADKIEDVTRQVACCKA 5154

RESULT 32
US-10-995-561-779
; Sequence 779, Application US/10995561

```

```

; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: C0001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 779
; LENGTH: 5415
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-995-561-779

Query Match      2.8%; Score 98; DB 6; Length 5415;
Best Local Similarity 18.9%; Pred. No. 1.2e+02;
Matches 154; Conservative 103; Mismatches 265; Indels 292; Gaps 43;

QY      9 AITIOKTYRRHQAARRMQR-----CNWQIF--ONTLYASBQ-- 43
DB      4456 ALLEQKW---HYVSSKMERKSKLEBALNATEFQNSLOEFINWLTLAGSINIASPPSL 4512

QY      44 -----DQAELYKFF--NDLIKMPQ-----AAGRNQYOGSAH-----YSV----- 77
DB      4513 ILMTVLSQIEBKHFANRVAHNDQIIEIDQNGNQLFLOKODVVLIRNLVSVQSRWE 4572

QY      78 -----LDDKDLVBERFGDIVNAKIE--LPIRKNHIDLLIDV----- 111
DB      4573 KVVQRSIERGRSLDDAKRAKQFHEAMKCLIDWLEDSHLSBELSISNDPKIKLOLSK 4632

QY      112 ---FRKGRNRLHPKYVALILREAAKSLQ--LPIISPVSTAVSOV-----TVCG--- 157
DB      4633 HKEFQKTLGSK-QPYVDITI--RTGRALKKTKLLPBDTQKDNFLGEVNDKMDTVCGSKV 4689

QY      158 DLHGKLDLLVVLHKNGLPSSSNPYFNGDFVDRGR-----GLE 197
DB      4690 ERQHKLEAL-----LFSGQFMDALQALVDMLYKVEPOLAEDQPVHGDD 4734

QY      198 VLLLLSLYLAFPNV-----FLNRGNHDSVNNAR-----YGFIREV----- 235
DB      4735 LVNMLDAHKVFOKEIGKRTGVQLKRGRE--LIENSRDQTTWVGQLOELSTRMDTVG 4793

QY      236 -----ESKYPRNKRIILAFIDEVY---RWLPLGSLVNSRVLLVHGFSSTSLDIKSI 286
DB      4794 KLSVSKQSLQALKAQAEFRDVTVMLEWLSA---EQTLRFRALPDDEA--LQSLI 4848

QY      287 DRGKYVSLRPLTDSRPDKTEWQOIFDIMMSDPQATWG-----CVPTILGAGVWFG 340
DB      4849 D-----THKEFMKVEKRV-DV---NSAVAMGEVILAVCHPDCITTIKHWI- 4891

QY      341 PDVTNDFLOHRRLSYVIRSHCEKPNGHEFMNDKIIITISASNYVALISGNKAYIYRLNQ 400
DB      4892 -----TIIRAREEVLTWAK-----QHQRLETLASE-----LVANAE 4924

QY      401 LMPHFVOYISAASQT-----KRLSFQRMGIYESSALKELAVRMRDHRD 444
DB      4925 LLEELLAMIQMAETTLIQDOEPRIQONIDRYKAL-----IAEHQTFMEBMTKQPDVD 4977

QY      445 ELBDEFK--YDPKDSGYT--SISHMCKYMENTYKGLP----- 479
DB      4978 RVTKTYKRNKIBPTHAPFIEKSRSGRKSISQPTPPMPLISQSEAKNPRINQLSARWQ 5037

QY      480 -WRLRLDKLAGTDSQKYNRKTDLDTDVITLEABD-----GMSVMDALYANKASLV 532
DB      5038 VMLLALER-----GRGLN-----DALDRLEELKEFANPDFDVARKKYMMAHAKSRVA 5086

QY      533 AIFNIIDADNSGKITLDEFTALID--LVANHPGAYSAEMLEKCRMDLNGDKVNDLNEF 591
DB      5087 DFFRIDKDDGKITRQEF--IDGILASKFP--TTLKMTAVADIDRDDGDYIDYEF 5441

```

```

QY      592 LBAFRLS-DLHRKEQODENIRRRSGRSPVATXA 624
DB      5142 VALLHPNKDAVPTTDADKIEDEVTROVAQCKCA 5175

RESULT 33
US-10-995-561-775
; Sequence 775; Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: C0001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 775
; LENGTH: 5464
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-995-561-775

Query Match      2.8%; Score 98; DB 6; Length 5464;
Best Local Similarity 18.9%; Pred. No. 1.2e+02;
Matches 154; Conservative 103; Mismatches 265; Indels 292; Gaps 43;

QY      9 AITIOKTYRRHQAARRMQR-----CNWQIF--ONTLYASBQ-- 43
DB      4456 ALLEQKW---HYVSSKMERKSKLEBALNATEFQNSLOEFINWLTLAGSINIASPPSL 4512

QY      44 -----DQAELYKFF--NDLIKMPQ-----AAGRNQYOGSAH-----YSV----- 77
DB      4513 ILMTVLSQIEBKHFANRVAHNDQIIEIDQNGNQLFLOKODVVLIRNLVSVQSRWE 4572

QY      78 -----LDDKDLVBERFGDIVNAKIE--LPIRKNHIDLLIDV----- 111
DB      4573 KVVQRSIERGRSLDDAKRAKQFHEAMKCLIDWLEDSHLSBELSISNDPKIKLOLSK 4632

QY      112 ---FRKGRNRLHPKYVALILREAAKSLQ--LPIISPVSTAVSOV-----TVCG--- 157
DB      4633 HKEFQKTLGSK-QPYVDITI--RTGRALKKTKLLPBDTQKDNFLGEVNDKMDTVCGSKV 4689

QY      158 DLHGKLDLLVVLHKNGLPSSSNPYFNGDFVDRGR-----GLE 197
DB      4690 ERQHKLEAL-----LFSGQFMDALQALVDMLYKVEPOLAEDQPVHGDD 4734

QY      198 VLLLLSLYLAFPNV-----FLNRGNHDSVNNAR-----YGFIREV----- 235
DB      4735 LVNMLDAHKVFOKEIGKRTGVQLKRGRE--LIENSRDQTTWVGQLOELSTRMDTVG 4793

QY      236 -----ESKYPRNKRIILAFIDEVY---RWLPLGSLVNSRVLLVHGFSSTSLDIKSI 286
DB      4794 KLSVSKQSLQALKAQAEFRDVTVMLEWLSA---EQTLRFRALPDDEA--LQSLI 4848

QY      287 DRGKYVSLRPLTDSRPDKTEWQOIFDIMMSDPQATWG-----CVPTILGAGVWFG 340
DB      4849 D-----THKEFMKVEKRV-DV---NSAVAMGEVILAVCHPDCITTIKHWI- 4891

QY      341 PDVTNDFLOHRRLSYVIRSHCEKPNGHEFMNDKIIITISASNYVALISGNKAYIYRLNQ 400
DB      4892 -----TIIRAREEVLTWAK-----QHQRLETLASE-----LVANAE 4924

QY      401 LMPHFVOYISAASQT-----KRLSFQRMGIYESSALKELAVRMRDHRD 444
DB      4925 LLEELLAMIQMAETTLIQDOEPRIQONIDRYKAL-----IAEHQTFMEBMTKQPDVD 4977

QY      445 ELBDEFK--YDPKDSGYT--SISHMCKYMENTYKGLP----- 479
DB      4978 RVTKTYKRNKIBPTHAPFIEKSRSGRKSISQPTPPMPLISQSEAKNPRINQLSARWQ 5037

```





Db 447 KGLSFV--BKCPSEHE 462

RESULT 36

US-10-131-826A-190

/ Sequence 190, Application US/10131826A

/ Publication No. US20050245730A1

/ GENERAL INFORMATION:

/ APPLICANT: Baker, Kevin P.

/ APPLICANT: Beresini, Maureen

/ APPLICANT: Defoyse, Laura

/ APPLICANT: Desnoyers, Luc

/ APPLICANT: Filvaroff, Ellen

/ APPLICANT: Gao, Wei-Qiang

/ APPLICANT: Gerltzen, Mary E.

/ APPLICANT: Goddard, Audrey

/ APPLICANT: Godowski, Paul J.

/ APPLICANT: Gurney, Austin L.

/ APPLICANT: Sherwood, Steven

/ APPLICANT: Smith, Victoria

/ APPLICANT: Stewart, Timothy A.

/ APPLICANT: Tumas, Daniel

/ APPLICANT: Watanabe, Colin K

/ APPLICANT: Wood, William

/ APPLICANT: Zhang, Zemin

/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

/ TITLE OF INVENTION: ACIDS ENCODING THE SAME

/ FILE REFERENCE: P330R1C128

/ CURRENT APPLICATION NUMBER: US/10/131,826A

/ PRIOR FILING DATE: 2002-04-24

/ PRIOR FILING DATE: 1997-06-18

/ PRIOR FILING DATE: 1997-08-26

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

/ PRIOR FILING DATE: 1997-09-17

Db 305 EVHYDNPTVE-----GLDNGSLR-LFTYM-----DIRKY---DAGVIBAG 342

QY 465 HWCKVMENTYKGLPWRLLRDLAPGDSQKVNRTLDLDTDTVLLEBADGMSVMDAL 524

Db 343 LWSLPHHT-----PQMPFQSGHCHTLECL-----EAL 373

QY 525 YANKASLVAFNIIADNSGEITLDEFETALDLVAHMGAYSKAMLEKCMMDLNG-D 583

Db 374 EABKSGHIVFAV-----LHNAHLAQRGIRLRHFRKGRKEMTLAYD 414

QY 584 GRVDLNEFLAFLRLSDLRKREQ 606

Db 415 DDFDFN--FOEFO-----YIKREQ 431

RESULT 37

US-10-118-590-2

/ Sequence 2, Application US/10118590

/ Publication No. US2005027761A1

/ GENERAL INFORMATION:

/ APPLICANT: KENNETH RHODES, MARIA BETTY, HUAI-PING LING, AND FRANK AN

/ TITLE OF INVENTION: POTASSIUM CHANNEL INTERACTORS AND USBS THEREFOR

/ FILE REFERENCE: MNI-070

/ CURRENT APPLICATION NUMBER: US/10/118,590

/ CURRENT FILING DATE: 2002-04-08

/ PRIOR APPLICATION NUMBER: US/09/298,731

/ PRIOR FILING DATE: 1999-04-23

/ NUMBER OF SEQ ID NOS: 47

/ SOFTWARE: PatentIn Ver. 2.0

/ SEQ ID NO 2

/ LENGTH: 216

/ TYPE: PRT

/ ORGANISM: Homo sapiens

US-10-118-590-2

Query Match 2.8%; Score 97; DB 6; Length 216;

Best local similarity 20.8%; Pred. No. 1.2;

Matches 50; Conservative 33; Mismatches 69; Indels 88; Gaps 9;

QY 424 MGIVSSALKEIYVNRDRDELDEFRKYDPRDSGYISISMKCKMVENTYKGLPWRLL 483

Db 5 MGFSSIAQTKQ---RRPKDKIIDE-----LEMTVCHRPRGLBQLEAGTNP 49

QY 484 RDKLAPGDSQKVNRTLDLDTDTVLLEBADGMSVMDALVA---NKASLVA---IFN 536

Db 50 KREL-----QVLRGFKNGECSGVNE-----DTFQIYAQFPFGDASTYAHYLRN 96

QY 537 IIDADNSGEITLDEFETALDLV-----AHMGAYSKAMLEKCR----- 576

Db 97 APDTGTGSKVRFEDFTALSLRGTVHEKLRMTFNLVDINKGVIKKEEMDIYKAIYD 156

QY 577 -----MMDLNGDGKVDLNEFLAFLRLSDLRKREQQDNIRK 612

Db 157 MMGKTYTVLKEDTPRQHVDFVQKDKKKGIVTLDEFLBSC-----QEDDNIRK 207

RESULT 38

US-10-118-590-6

/ Sequence 6, Application US/10118590

/ Publication No. US2005027761A1

/ GENERAL INFORMATION:

/ APPLICANT: KENNETH RHODES, MARIA BETTY, HUAI-PING LING, AND FRANK AN

/ TITLE OF INVENTION: POTASSIUM CHANNEL INTERACTORS AND USBS THEREFOR

/ FILE REFERENCE: MNI-070

/ CURRENT APPLICATION NUMBER: US/10/118,590

/ CURRENT FILING DATE: 2002-04-08

/ PRIOR APPLICATION NUMBER: US/09/298,731

/ PRIOR FILING DATE: 1999-04-23

/ NUMBER OF SEQ ID NOS: 47

/ SOFTWARE: PatentIn Ver. 2.0

/ SEQ ID NO 6

/ LENGTH: 216

/ TYPE: PRT

/ ORGANISM: Homo sapiens

US-10-118-590-6

Query Match 2.8%; Score 97; DB 6; Length 216;

Best local similarity 20.8%; Pred. No. 1.2;

Matches 50; Conservative 33; Mismatches 69; Indels 88; Gaps 9;

QY 424 MGIVSSALKEIYVNRDRDELDEFRKYDPRDSGYISISMKCKMVENTYKGLPWRLL 483

Db 5 MGFSSIAQTKQ---RRPKDKIIDE-----LEMTVCHRPRGLBQLEAGTNP 49

QY 484 RDKLAPGDSQKVNRTLDLDTDTVLLEBADGMSVMDALVA---NKASLVA---IFN 536

Db 50 KREL-----QVLRGFKNGECSGVNE-----DTFQIYAQFPFGDASTYAHYLRN 96

QY 537 IIDADNSGEITLDEFETALDLV-----AHMGAYSKAMLEKCR----- 576

Db 97 APDTGTGSKVRFEDFTALSLRGTVHEKLRMTFNLVDINKGVIKKEEMDIYKAIYD 156

QY 577 -----MMDLNGDGKVDLNEFLAFLRLSDLRKREQQDNIRK 612

Db 157 MMGKTYTVLKEDTPRQHVDFVQKDKKKGIVTLDEFLBSC-----QEDDNIRK 207

RESULT 38

US-10-118-590-6

/ Sequence 6, Application US/10118590

/ Publication No. US2005027761A1

/ GENERAL INFORMATION:

/ APPLICANT: KENNETH RHODES, MARIA BETTY, HUAI-PING LING, AND FRANK AN

/ TITLE OF INVENTION: POTASSIUM CHANNEL INTERACTORS AND USBS THEREFOR

/ FILE REFERENCE: MNI-070

/ CURRENT APPLICATION NUMBER: US/10/118,590

/ CURRENT FILING DATE: 2002-04-08

/ PRIOR APPLICATION NUMBER: US/09/298,731

/ PRIOR FILING DATE: 1999-04-23

/ NUMBER OF SEQ ID NOS: 47

/ SOFTWARE: PatentIn Ver. 2.0

/ SEQ ID NO 6

/ LENGTH: 216

/ TYPE: PRT

/ ORGANISM: Homo sapiens

US-10-118-590-6

Query Match 2.8%; Score 97; DB 6; Length 216;

Best local similarity 20.8%; Pred. No. 1.2;

Matches 50; Conservative 33; Mismatches 69; Indels 88; Gaps 9;

```
; ORGANISM: Mus musculus
US-10-118-590-6

Query Match
Best Local Similarity 2.8%; Score 97; DB 6; Length 216;
Matches 50; Conservative 33; Mismatches 69; Indels 88; Gaps 9;

QY 424 MGIVSSALKELAVRMHRDELDEFKVPKOSGIYSISHCKWMENTYKGLPRL 483
DB 5 MGRSSSLQTKQ-----RRPSKQIBDE-----LEMTVCHREPGLEQLAQNFT 49
QY 484 RDKLAPGDSQKVVNRTLDLDTDVILEADGWSVDALYA-----NKASLYA--IFN 536
DB 50 KREL-----QVLYRGKACPSGVNHE-----ETKQIYAQPPFGDASTYAHYFN 96
QY 537 IIDANGSEITLDEFETAILLV-----AMPGAYSXAMLEKCR--- 576
DB 97 AFDTTQTSVAFEDVETALSLKQTVHEKLRMTFNLYDKDKGYINKKEMWDIVKAIYD 156
QY 577 -----MDLNGDKTVDFLEAFRLSLDLHREKQDENIR 612
DB 157 MNGKTYPVAKEDTPRQHDVDFQKMDKXKGIYTLDEFLSC-----QBDNIMR 207

RESULT 39
US-10-878-556A-161
; Sequence 161, Application US/10878556A
; Publication No. US20050266399a1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann La-Roche Inc.
; TITLE OF INVENTION: HCV regulated protein expression
; FILE REFERENCE: 21762
; CURRENT APPLICATION NUMBER: US/0/878.556A
; CURRENT FILING DATE: 2004-06-28
; NUMBER OF SEQ ID NOS: 199
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 161
; LENGTH: 814
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: sw_hum/f1f39_human
; DATABASE ENTRY DATE: 1997-11-01
US-10-878-556A-161

Query Match
Best Local Similarity 2.8%; Score 97; DB 6; Length 814;
Matches 73; Conservative 53; Mismatches 119; Indels 124; Gaps 16;

QY 185 NGDF---VDRGKGLVLLLLSLYLAFPNAVFLNRGNHDSVMNARYGPIREVSCTP 240
DB 522 NGDYLCVVDRTPKGTQGVVT-----NPEIFMRREKQVVDVEMK-- 562
QY 241 RNHRRIIAFIDEVYRMPLGSLVLSRVLYHG-----GPSDSLSLLKSIDRGK 290
DB 563 ---ETTLAF-----AMEPNG---SKFVLHGEAPRISVSFYHYKXNKGIKELIKMPDK-- 608
QY 291 YVSLIRPLTDGBELDKTEMQOIFDIMWSDPOATMGCVNPLRAGVFGPDVTNFIQR 350
DB 609 -----QANTIFWS-PQCGFVVLVLAGLRSMNGALAFVDTSDCYMN 647
QY 351 HRLSVIRSHCKENGHEFMNDKITYFS-----ASNYVAIGSNKGAIVIRLANQLMHP 405
DB 648 IAEHYMASDVEMDPLG-----RYVTVSVSWSHKVDNAVYMLTFQGRLLQKNK--DRF 699
QY 406 VQY-----ISASQTKRLS-----FKQRMGIVESSALKELAVRMRDRHDELD 448
DB 700 COLMRPPPTLLSQEQIKQIKDKLKKYSKIFEQQRISQSSKASKELEVERR---TME 755
QY 449 EFRKIDPKDSGISISHCKWMENTYKGLPRLRLDKLAPGDSQKVVN-----RTL 502
DB 756 DFRKY-----RKMAQELYMKGKNERLRLRGVDTDLSDSVDDMBEETI 799
```

```
QY 503 DLDPTVIL 511
DB 800 EFPVTEETI 808

RESULT 40
US-11-115-639-50
; Sequence 50, Application US/11115639
; Publication No. US20050282242A1
; GENERAL INFORMATION:
; APPLICANT: Rotherstein, David
; APPLICANT: Murphy, Christopher
; APPLICANT: MacNeil, Ian
; TITLE OF INVENTION: SCREENING ASSAYS FOR ANTIMICROBIAL
; FILE REFERENCE: 50150/075003
; CURRENT APPLICATION NUMBER: US/11/115,639
; CURRENT FILING DATE: 2005-04-27
; PRIOR APPLICATION NUMBER: 60/566,858
; PRIOR FILING DATE: 2004-04-30
; PRIOR APPLICATION NUMBER: 60/565,679
; PRIOR FILING DATE: 2004-04-27
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 1184
; TYPE: PRT
; ORGANISM: L. monocytogenes
US-11-115-639-50
```

```
Query Match
Best Local Similarity 2.8%; Score 97; DB 7; Length 1184;
Matches 157; Conservative 88; Mismatches 257; Indels 252; Gaps 41;

QY 16 YRRHOAREMQRRCNQIIPONLEVAS--EODQAEIYFNDLTKHMPQAGRKQYOGSA 73
DB 11 YGRHRTKRSFAR-----ISEVLELPNILEIQTASVCMFDEGLAREMRDPSIEDPAGNL 65
QY 74 HVSVLDDKDDLVEBFQDIVNAKIELPIRKNHIDLLIDVFRKKGKGNRLHPKYVALILREA 133
DB 66 SLERID-----YDLGEPKYSVEESKN-----RDNVAAVPLRVKRL----- 101
QY 134 KSLKQLPNISPVSTAISQVTVCGD--LHGKLDLVLVYHKG-----LPSSNPVY 183
DB 102 -----INKETGEVVDQEVFMGDPFLMTEMGFTI---NGAERYIVSQLVRSPEVY- 148
QY 184 FNGDPVDRGKRGLEVLILLSLYLAFPNAVFLNRG-----NHEDSVNAR----- 228
DB 149 FNGKLDKNGKKG-----FQSTVIYPRNGAMLEYETDAKDVAHVIRIDRTKLPV 195
QY 229 -----YGFIREV-----ESKYPN-----HKRIIAFIDEVYRMPLG--SV 262
DB 196 TVLLRALGFGSDQDEIIDLIGNDNYLNTLEKMDTNAEKALL-----EYERLRGGEPTV 251
QY 263 LNSKRVLYHGGFSDSTSLDIKSIDRK-----YVSLIRPLTD----- 301
DB 252 DNARSILV--SRFPDPKRYDL--ASVGRYKINKKHLKNRLFNQTLAETLVPEPTEGELIASK 309
QY 302 GEPIDKTEMQOIFDIMWSDPOATMGCVNPLRAGV--WFGPDVTNFIQSRHSVYIRSH 360
DB 310 GDILDRANDLQI-----IPNLENGVGFPTTLR--TDGVMESVLVQSIKIY 353
QY 361 ECKPNHGEFMNDKITYIFSASNYVAIGSNKGAIVIRLANQLMHPVQYISAASQTKRLS- 419
DB 354 --APNDE-----KEINI-----IGN--AYIBEN-----VGHITPSDIISSISY 368
QY 420 -FKQRMGIVESSALKELA--VRMDHDELDEDERKVPKOSGIYSISHCKV-----ME 471
DB 389 FPNLHGVGDITDIDLGNRLRSVGBLLQNGR-----IGLSRMRVVRBERMSIQ 439
QY 472 NNTKLGAPRL-----LRDKLAPGDSQKVVNRTLDLDTDVILEA-----E 514
DB 440 DMTTI--TPQULINIRPVVASIKKEPFGSSQLSQFMNQINPILDELITYKRLSALRGGLTRE 498
```

```

Qy 515 ADGMSVMDALYANKASU-----VAIFNIIDADNSGEITLDER-----E 552
Db 499 RAGYFRADVHYSHYGRMCPIETPEGNIGLFTYL---SSPAKVNKGPIETPYRRVDP 554
Qy 553 T-----AIDLVAHMPGAYSKAMLEKCRMDLNGDKVDLNEFLAFLSLDIHRKQOD 607
Db 555 TNRVTDKIDYLTADAEDNYYVVAQANSK-----LDEQOTFTEBEVMARFSENL---AVEK 606
Qy 608 ENIRRRSTGRPSYAKTATDPVTLLADKISNTLY 641
Db 607 ERIDYMDVSPKQYVSVATACIPLENDSDNRALM 640

```

Search completed: January 20, 2006, 20:04:22  
 Job time : 37 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

## OM protein - protein search, using SW model

Run on: January 20, 2006, 19:39:22 ; Search time 40 Seconds

(without alignments)  
1589.982 Million cell updates/sec

Title: US-09-463-733-1

Perfect score: 3442  
Sequence: 1 MDENAIRAIFQKXTRRHO.....VEHDIDPTDCSKVIDPKKS 661Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summariesDatabase : PIR 80:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3442	100.0	661	1 A42287	phosphoprotein pho
2	1317	38.3	707	2 T42239	probable phosphor
3	1305	37.9	722	2 T34072	hypothetical prote
4	593.5	17.2	704	2 T14614	hypothetical prote
5	569	16.5	499	1 A55346	phosphoprotein pho
6	567.5	16.5	498	1 S52370	phosphoprotein pho
7	558	16.2	526	1 T45058	phosphoprotein pho
8	557.5	16.2	533	2 B84858	phosphoprotein pho
9	541	15.7	479	1 T46576	phosphoprotein pho
10	539	15.1	473	1 T40391	phosphoprotein pho
11	479.5	13.9	513	1 S52371	phosphoprotein pho
12	444	12.9	413	2 T51611	phosphoprotein pho
13	441	12.8	326	2 T09995	phosphoprotein pho
14	436.5	12.7	314	2 S52371	phosphoprotein pho
15	435.5	12.7	327	2 A32550	phosphoprotein pho
16	434	12.6	729	2 T09913	probable phosphor
17	433.5	12.6	326	2 T09544	phosphoprotein pho
18	429.5	12.5	323	2 A32349	phosphoprotein pho
19	429	12.5	322	2 B32550	phosphoprotein pho
20	428	12.4	318	2 S20882	phosphoprotein pho
21	426.5	12.4	317	2 T03594	phosphoprotein pho
22	425	12.3	401	1 T21288	phosphoprotein pho
23	424.5	12.3	312	2 S31086	phosphoprotein pho
24	421	12.2	323	2 T09550	phosphoprotein pho
25	421	12.2	326	2 T09548	phosphoprotein pho
26	420.5	12.2	317	2 T03304	probable phosphor
27	420	12.2	291	1 T21322	phosphoprotein pho
28	419.5	12.2	321	2 S31088	phosphoprotein pho
29	418.5	12.2	310	1 T25993	phosphoprotein pho

30	418.5	12.2	312	2 S32595	phosphoprotein pho
31	418.5	12.2	312	2 S24264	phosphoprotein pho
32	418.5	12.2	325	2 T09547	phosphoprotein pho
33	413.5	12.0	401	2 T16936	phosphoprotein pho
34	413	12.0	302	1 PAFPIA	phosphoprotein pho
35	412	12.0	316	2 S26225	phosphoprotein pho
36	411.5	12.0	308	1 PABY3	phosphoprotein pho
37	411.5	12.0	333	1 T16476	phosphoprotein pho
38	409.5	11.9	382	1 T19701	phosphoprotein pho
39	409	11.9	316	2 S29317	phosphoprotein pho
40	408	11.9	312	2 S31089	phosphoprotein pho
41	408	11.9	322	1 S31087	phosphoprotein pho
42	407.5	11.8	323	2 T76572	phosphoprotein pho
43	407.5	11.8	337	1 C32550	phosphoprotein pho
44	406.5	11.8	337	1 T76573	phosphoprotein pho
45	405.5	11.8	323	1 S35699	phosphoprotein pho

## ALIGNMENTS

```
RESULT 1
A42287
phosphoprotein phosphatase (BC 3.1.3.16) rdgc - fruit fly (Drosophila melanogaster)
N/Alternate names: retinal degeneration protein C
C/Species: Drosophila melanogaster
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 12-Jul-2004
C/Accession: A42287, S27811
R/Steale, F. R.; Washburn, T.; Rieger, R.; O'Touse, J. B.
Cell 69, 669-676, 1992
A/Title: Drosophila retinal degeneration C (rdgc) encodes a novel serine/threonine prot
A/Reference number: A42287; MUID:9226398; PMID:1316807
A/Accession: A42287
A/Molecule type: DNA
A/Residues: 1-661 <STB>
A/Cross-references: UNIPROT:P40421, UNIPARC:UPI000000728A; EMBL:W89628; NID:G158237; PII
A/Note: sequence extracted from NCBI backbone (NCBIN:103017, NCBIN:103023, NCBIN:103028,
, NCBIN:103104, NCBIPI:103001)
C/Genetics:
A/Gene: rdgc
A/Cross-references: FlyBase:FBgn0004366
A/Introns: 5/1; 47/3; 74/3; 116/3; 172/3; 240/3; 290/3; 312/3; 374/3; 510/3; 544/1
C/Function:
A/Description: catalyzes the hydrolytic dephosphorylation of protein-phosphoserine and p
A/Note: prevents light-induced retinal degeneration
C/Superfamily: serine/threonine protein phosphatase with BF-hands; calmodulin repeat hom
C/Keywords: calcium binding; BF hand; iron; metalloprotein; nucleus; phosphoric monoeste
F:132-221/Domain: phosphoprotein phosphatase homology <PFC>
F:152-221/Domain: phosphatase core homology <BPT>
F:1526-558/Domain: calmodulin repeat homology <BPT>
F:1566-598/Domain: calmodulin repeat homology <BPT>
F:158,160,187/Binding site: iron (Aep, His, Asp) #status predicted
F:187,219,271,360/Binding site: zinc (Aep, Asp, His, His) #status predicted
F:190,220,384/Active site: Asp, His, Tyr #status predicted
F:191,333/Binding site: substrate phosphate (Arg) #status predicted

Query Match 100.0%; Score 3442; DB 1; Length 661;
Best Local Similarity 100.0%; Pred. No. 1.5e-208;
Matches 661; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDENAIRAIFQKXTRRHOAREMQRRCNMOIFONLBYASQDAELYKFPNDLIKMP 60
DB 1 MDENAIRAIFQKXTRRHOAREMQRRCNMOIFONLBYASQDAELYKFPNDLIKMP 60
QY 61 QAAGRKNQYQSAHVSVDKDDLYVEFGDIYNAKIELEPIRKNHIDLDIVRKKRGNTL 120
DB 61 QAAGRKNQYQSAHVSVDKDDLYVEFGDIYNAKIELEPIRKNHIDLDIVRKKRGNTL 120
QY 121 HKRYVALILREAKSLKQLEPNISPVSTAVSQOVTGDLHGLDLDLVLLHNGKLPSSSN 180
DB 121 HKRYVALILREAKSLKQLEPNISPVSTAVSQOVTGDLHGLDLDLVLLHNGKLPSSSN 180
QY 181 PYVENGDFVRDGRGLGVLLILSLYLAEPNAVPLNRGNHEDSVNNAARYGPIREVSXTP 240
```

```

Db      181 PVTNMGDFDRKRGKGVLLLLSLYLAFPAVFLNRGNHDSVMNAKGFIREVSKTP 240
Qy      241 RNHKKILAFIDEVYRNLPVGSVLSNRVLIVHGFSDSTSLDIKSIDRGKVSILRPPLT 300
Db      241 RNHKKILAFIDEVYRNLPVGSVLSNRVLIVHGFSDSTSLDIKSIDRGKVSILRPPLT 300
Qy      301 DDEPDLKTEWQOIFDIWMSDPOATMGCVPTLRGAGVWGPVPTNPLQRHLSVYISHS 360
Db      301 DDEPDLKTEWQOIFDIWMSDPOATMGCVPTLRGAGVWGPVPTNPLQRHLSVYISHS 360
Qy      361 ECKPNGHEFMHDKITITISASNYAIGSNKGAAYIRLNQMLPHFYQVYISAASQTKRLSF 420
Db      361 ECKPNGHEFMHDKITITISASNYAIGSNKGAAYIRLNQMLPHFYQVYISAASQTKRLSF 420
Qy      421 KQRMGIVBSSALKEIAYNRMRDRDELDEFKRYDPSGYSISHWCKMENVTKLGLPW 480
Db      421 KQRMGIVBSSALKEIAYNRMRDRDELDEFKRYDPSGYSISHWCKMENVTKLGLPW 480
Qy      481 RLRLKTLAQTDSOKVNVARTLDDLTQVYILAEADGMSVMALYANKASLVAIFNIIDA 540
Db      481 RLRLKTLAQTDSOKVNVARTLDDLTQVYILAEADGMSVMALYANKASLVAIFNIIDA 540
Qy      541 DMSGITLDEFETALDILVAHMPGAYSKAEMLEKCMMDLNGDKVDLNEFLAEARLSDL 600
Db      541 DMSGITLDEFETALDILVAHMPGAYSKAEMLEKCMMDLNGDKVDLNEFLAEARLSDL 600
Qy      601 HKKQODENIRRRSTGRPSVAATADPTVLADKISKNTLVVEHDIDPTDCSKVIDPPK 660
Db      601 HKKQODENIRRRSTGRPSVAATADPTVLADKISKNTLVVEHDIDPTDCSKVIDPPK 660
Qy      661 S 661
Db      661 S 661

```

## RESULT 2

T42239  
 Probable phosphoprotein phosphatase (BC 3.1.3.16) - *Caenorhabditis elegans*  
 C/Species: *Caenorhabditis elegans*  
 C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 12-Jul-2004  
 C/Accession: T42239  
 R:Sherman, P. M.; Sun, H.; Macle, J. P.; Williams, J.; Smallwood, P. M.; Nathans, J.  
 Proc. Natl. Acad. Sci. U.S.A. 94, 11639-11644, 1997  
 A/Title: Identification and characterization of a conserved family of protein serine/thr  
 A/Reference number: Z22115; MUID:97471020; PMID:9326663  
 A/Accession: T42239  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-707 <SHE>  
 A/Cross-references: UNIPROT:O01921, UNIPARC:UP10000060867, EMBL:AF023454, NID:92586408,  
 C/Superfamily: serine/threonine protein phosphatase with EF-hand; calmodulin repeat hom  
 C/Keywords: EF hand; phosphoric monoester hydrolase; serine/threonine-specific phosphatase

```

Query Match      38.3%; Score 1317; DB 2; Length 707;
Best Local Similarity 42.1%; Pred. No. 5.3e-75;
Matches 266; Conservative 116; Mismatches 200; Indels 50; Gaps 12;

Qy      6 IBAALFIQMYRRHQARRMQRCKNQIFONTLEYASEDOAEIYKFFNDLIKHPQOAGR 65
Db      83 IKSALILQKMYRRCARLEARRATWQIFALAEAGEOQLKRYVWITLYDFADVIR 142
Qy      66 KNOYQG--SAHVSLVD--KDLVEEFGDIVNAKIF-----LPIRK 102
Db      143 GGVENGRNSPLMSALSHYAKPSLMSRGFTVKMLDPTVWDIDIRNYKGPFLSLPLDK 202
Qy      103 NHIDILIDVFRKKRGRRLAPKYVALILREAAKSLKQLPNIISPVSTAVSQOVTAVCGDILNK 162
Db      203 PQVANKITEAF--KVNKVLHPKYVLMILHARAKIFKAMPVSYSRISISQVITICDILNK 260
Qy      163 LDLLVLVLRKNGILPSSSNPYVFNQDFVDRGKRGLEVLALLSLYLAFPAVFLNRGNHSD 222
Db      261 PDDLCLILVLRKNGIPSVNPIYIFNGDFVDRGGGSIIVLCVLPALVIVDPMSIYLNRGNHSD 320

```

```

Qy      223 SVMNARVGFIREVSKTPRNHKKILAFIDEVYRNLPVGSVLSNRVLIVHGFSDSTSLDI 282
Db      321 HINMLRFGFIRELSTKYKDLSTPIRLLEDFVSMPLATIIDRDI FVHGQISDQTEVSK 380
Qy      283 IKSIDRGKVSILRPPLLDGRPLDK-----TEMQOIFDIWMSDPOATMGCVPTLRGA 335
Db      381 LDKTFRHRFQSVLPVPVKKGNSEKENSANVNDERKQMLDIWMSDPOKQKCGMPVFRGG 440
Qy      336 GWFPGPDVTDNPLQRHLSVYIRSHCKPNGBPMHDKITITISASNYAIGSNKGAAYI 395
Db      441 GSYFADITAFLEHGFRLLVVRGHECKFEYEFSSHNTCLTVFSASNYIFETGSRGAYV 500
Qy      396 R-LNQMLPHFYQVYISAASQTKRLS-FQRMGIVBSSALKEIAYNRMRDRDELDEFKRY 453
Db      501 KFIGSKQPHFYQVY--ASKTKRSTYLERGIVBSSAVKELKELSGFHTDLOKRFBIW 558
Qy      454 DPKSGYISISHMCKMENVTKLGLPWRLLRDKLAP--QTDGOKVNVARTLDDLTQVILE 512
Db      559 DIERSGKPLIKMSDQVRRITGLNLPWIALAPKVAITLSBDQKYVYKEDRAIAQVGTHA 618
Qy      513 AEADGMSVMALYANKASLVAIFNIIDMSGITLDEFETALDILVAHMPGAYSKAEM- 571
Db      619 QEKD---IVBSLYHRKSTLETLPFRMDKONGQVSMKEFIDACEVL-----GKTKRPLQ 670
Qy      572 ----LEKCRMDLNGDKVDLNEFLAEARLSD 599
Db      671 TDYISQIAESIDFNKDFIDNELLEAFRLVD 702

```

## RESULT 3

T34072  
 Hypothetical protein F23H11.8 - *Caenorhabditis elegans*  
 C/Species: *Caenorhabditis elegans*  
 C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 12-Jul-2004  
 C/Accession: T34072  
 R:Roehlfing, T.; Wohldmann, P.  
 submitted to the EMBL Data Library, May 1997  
 A/Description: The sequence of C. elegans cosmid F23H11.  
 A/Reference number: Z21470  
 A/Accession: T34072  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-722 <ROH>  
 A/Cross-references: UNIPROT:O01921, UNIPARC:UP10000175983, EMBL:AF003389, PIDN:AACT1139.  
 A/Experimental source: strain Bristol N2; clone F23H11  
 C/Genetics:  
 A/Gene: CESP:F23H11.8  
 A/Map position: 3  
 A/Intons: 91/1, 139/3, 353/3, 579/1, 674/3  
 C/Superfamily: serine/threonine protein phosphatase with EF-hand; calmodulin repeat hom  
 C/Keywords: EF hand

```

Query Match      37.9%; Score 1305; DB 2; Length 722;
Best Local Similarity 41.6%; Pred. No. 3.1e-74;
Matches 266; Conservative 117; Mismatches 201; Indels 56; Gaps 13;

Qy      4 NAIRALFIQMYRRHQARRMQRCKNQIFONTLEYASEDOAE-----LYKFFNDLIK 57
Db      90 STIKSALILQKMYRRCARLEARRATWQIFALAEAGEOQLKRYVWITLYDFADVIR 149
Qy      58 HMPQAGRKNOYQG--SAHVSLVD--KDLVEEFGDIVNAKIF----- 97
Db      150 AMAENRGQGVNGNSPLMSALSHYAKPSLMSRGFTVKMLDPTVWDIDIRNYKGP 209
Qy      98 ---LPIRKNHIDILIDVFRKKRGRRLAPKYVALILREAAKSLKQLPNIISPVSTAVSQOVT 154
Db      210 TLSLPLDPQVANKITEAF--KVNKVLHPKYVLMILHARAKIFKAMPVSYSRISISQVIT 267
Qy      155 VCGDLHGRKLDLLVLVLRKNGILPSSSNPYVFNQDFVDRGKRGLEVLALLSLYLAFPAVFLNR 214
Db      268 ICGDLHGRKFDLCLILVLRKNGIPSVNPIYIFNGDFVDRGGGSIIVLCVLPALVIVDPMSIY 327

```

```

OY 215 INRGHESUNNARVGPIREVSCKPRNHKILAFIDBYVMPLGSLNARVLVHGF 274
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 328 LNRGHEHIMNLRGPIKELSTYKOLSTPIRLLBDFVSWLPATIIDIDIPVHGI 387
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 275 SDSTSLDIKSIDRGKYSILRPLPTDGEPIDK-----TEWQIPIIMSDPOATWC 327
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 368 SDQTEVSLDKIPIRHRFOSVLRFPYKMGESKEKSNANVDEMKIMDIMSDKONKC 447
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 328 VPTILGAGVFGPDVDTNPLQRHRLSYVIRSHCEKPGHBFMDNKLITTFASANTYAI 387
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 448 WPNVFRGGSGFYGADITASFLEKHGFRLLVRSHECKFGEFYSHNNTCLTVFASANYET 507
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 368 GSNKAYIR-LINOLMBHFVUYISAASOTKLS-FKOMGIVESSALKELAVRRDRDE 445
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 508 GSNRAAYVFKSGSKQPHFVOYM-ASKTHRKSTIRELVGESAVDELKELKSPFDT 565
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 446 LEDBERKDPKDSGYISISHCKWMENTKJGLPRLRLDRLAP-GTDSOKVNTNRTLDL 504
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 566 LQKEBEIMDIRSGCLPIKMSDCYERITGLNLPILALAPKATLSDEGKVMYKEBRI 625
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 505 LDTDVILEADGMSVMDALYANKRASLVAFINIIDADNSGEITLDEBETALDILVAHMPG 564
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 626 AQVGTHAQEKD--IVESLYVRHKSTLETLFRPMDKONGGVSKMEFIDACEVL-----G 677
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 565 AYSKXEM-----LEKCRMDLNGDGKVDLNEFLAFLUSD 599
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 678 KYTKRPLQTDYISQJAEISIDFNKQGFIDNLLELLEFRLLVD 717
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

```

RESULT 4
T14614
hypothetical protein - Trypanosoma cruzi
C:Species: Trypanosoma cruzi
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T14614
R:Anderson, B.; Aalund, L.; Pettersson, U.
submitted to the EMBL Data Library, March 1998
A:Description: 93.4 kb of complete sequence from chromosome 3 of Trypanosoma cruzi.
A:Reference number: Z18159
A:Accession: T14614
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-704 <AND>
A:Cross-references: UNIPARC:UPI000017B5B6; EMBL:AF052832; NID:g3063540; PID:g3063551; PI
C:Genetics:
A:Map position: 3

Query Match          17.2%; Score 593.5; DB 2; Length 704;
Best Local Similarity 33.0%; Pred. No. 1.5e-29;
Matches 173; Conservative 73; Mismatches 204; Indels 75; Gaps 15;

QY      128 ILREAKSLKQLPN-----SPVSTAVS-----QQVYVCGDLHGKLDLLVYHNRGLP 176
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      174 VFTDMGSLTNPENVRLSPVGARVSNGRINQSKVYVVDLHGQLADLLHILKEGMP 233
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      177 SSSNPYVNGDFVDRGKRGLVLLLSLYLAFPNAVFLNRGNHSDSYMANARYGPIREVE 236
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      224 NEGTYIIFNGDFVDRGANGVEVLLILFSLMACPKYVTLNRGNHCDYNDYGFVDEVS 293
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      237 SKYPRNHRILAFIDEVVRWPLPGSVLNSKRLIYHGGF--SDSTSLDLSIDRGKTVSI 294
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      294 TKYDNN---VRLVORCFCALPLATIIIGKKVFFVHGGPLPRKGVNIIDISRIQRFROI-- 348
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      295 LRPPLTDEBPLDKTEMQOIF-DIMWSDFOATMGCVPTLNRGAGVFGFDVYDNFLQRHRL 353
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      349 ---PMPD---YSQPEDEBIFQDLLMSDVEDVLDQGWRSRPRGAGVFGADVLTQERLONNGL 402
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      354 SYVIRSHCKCKNGRHFMDNKIITFFSANTYALGSKNGAL-ILNNQMLMHPVQIYSAA 412
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      403 ELVINSHEBCLRGHEHHDGKLLTVFSASNYDGPETNGSPAVVVGNDPESFHTYGVAE 462
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      413 SQT-----RLSEKMGVIVESA-----LKEILAVRRDRHREDELED 448
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db      443 DDVEVQITVLDEGETTPPLTGRVSSPATTSQSGLILRRARDDVLAVERITYQRRHLLA 522

Qy      449 BFRKYDPKDSGYISISHNCKWMENTXKGLPWRILRDKLAPGDSQKVANNRTL----- 502
Db      533 YPAKDDRIRKGSVMKLEWVEANRNVLTNDLPFPFLRGVLVADENRTVRWYSHFLVKFNHF 582

Qy      503 --DLIDTVILLEARDGMSVMDALYANKASLVAFNITIDADNSGEITLDEFETAIDLVA 560
Db      583 FOPLTANMIOSA---CHLTQQOQRANRSQY---VAKAFKQGVSVNFECSVIRAI DY 635

Qy      581 HMPGAYSKAEMLKCRMDLNDGKVDLNEF-----LEAPFLSD 599
Db      636 TM-----SEAQLFOLFVYLDEGGTGHIIDPKFNNMLSEMAAAYPLD 676

```

```

RESULT 5
A55346
phosphoprotein phosphatase (EC 3.1.3.16) PPT [validated] - rat
N.Alternate names: serine/threonine phosphatase pps
C.Species: Rattus norvegicus (Norway rat)
C.Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 28-Apr-2003
C.Accession: A55346
R.Beckert, W.; Kentrup, H.; Klump, S.; Schultze, J.E.; Joost, H.G.
J. Biol. Chem. 269, 22586-22592, 1994
A.Title: Molecular cloning of a protein serine/threonine phosphatase containing a putative
A.Reference number: A55346; MUID:94357899; PMID:80772208
A.Accession: A55346
A.Molecule type: mRNA
A.Residues: 1-499 <BEC>
A.Cross-references: UNIPARC:UPI0000172827; GB:X77237
A.Note: authors translated the codon AAG for residue 53 as Gln, and GTA for residue 496
C.Superfamily: protein phosphatase 5; phosphoesterase core homology; phosphoprotein pho-
C.Keywords: iron; metalloprotein; nucleus; phosphoric monoester hydrolase; zinc
F:28-61/Domain: tetratricopeptide repeat homology <TR1>
F:62-95/Domain: tetratricopeptide repeat homology <TR2>
F:96-129/Domain: tetratricopeptide repeat homology <TR3>
F:204-468/Domain: phosphoprotein phosphatase homology <PPP>
F:236-305/Domain: phosphoesterase core homology <PEC>
F:242,244,271/Binding site: iron (Asp, His, Asp) #status predicted
F:271,303,352,427/Binding site: zinc (Asp, Asn, His, His) #status predicted
F:274,304,451/Active site: Asp, His, Tyr #status predicted
F:275,400/Binding site: substrate phosphate (Arg) #status predicted

Query Match 16.5%; Score 569; DB 1; Length 499;
Best Local Similarity 33.7%; Pred. NO. 3.3e-26;
Matches 143; Conservative 73; Mismatches 140; Indels 68; Gaps 16;

11 FICQWYRRHQARRRMRQRCQWQIIFQNLRYVASQDQALYKFPNDLIKMPQAAGRKNQYQ 70
   :::::
95 YIKQYVYRRAASNNALGK-----FR-----AALND-----YETVAVKRNDDADARKYQ 137
   :::::

71 -----GSAN-VSYLDDKD---DLVEFG-DIVNAKIELPIRKNHIDL 107
   :::::

138 ECSKIVQKAPERAIADEHRRSVSDLSIESMTIEDYEGSKPLEDKVITFEWKD---- 193
   :::::

108 LIDVFRKKRGGRNRLAPKRVALLIREAAKSLQLPRISSVSTRANVQGYVPCDGLGKLDLL 167
   :::::
194 LMQWYKQOK-KLARRKAVQILVQVKEVLCKSLVETTLKEKEKIVQCGDTHGQFYDLL 251
   :::::

168 VLAHKNGLPSSSNPYVENQDFVDRGKGLFVLLLLSLYLAPPAVPLANGNHDSVANA 227
   :::::
252 NIFELNGLPSTNTYIFNGDFVDRGSSVAVILLRPFKLLYRDPHFLRLKGNHETDMNQ 311
   :::::

228 RYGFIREVSKYPRNHRKILAFIDEVYRWLPGLSVLNSRVLLVHGF--SDSTSLDIKS 285
   :::::
312 IYGEQGVAKAYT--AQWYELFESEVEMVPLAQCIINGKGLIMHGGLFSESDGVTLLDIRK 368
   :::::

286 IDRKQVYSILRPPLTQDBPLDKTEWQOIFPIMMSDPATMGCPNTRLRGSGWAFGPPVTD 345
   :::::
369 IERRN-----QEP--DSGPM-----CLLMSDPPQNGRSVSK-RGVSCQFGPDVTK 412
   :::::

346 NFIQRHLSVYIRSHCKPQNGHEFMNDKIIITFSASNYVAIGSNKCAVYRL--NNQIMPH 404
   :::::

```

Db 413 AFLEBNQDLIIRSHVKAQGYVAHGRCTVFSAPNYCDQGNKASYIHLQSGDLRQ 472

QY 405 FVOY 408

Db 473 FHOY 476

RESULT 6

SS2570

phosphoprotein phosphatase (EC 3.1.3.16) 5 [validated] - human

N/Alternate names: serine/threonine phosphatase pps

C/Species: Homo sapiens (man)

C/Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004

C/Accession: SS2570; PC4136

R/Chen, M.X.; McPartlin, A.E.; Brown, L.; Chen, Y.H.; Barker, H.M.; Cohen, P.T.W.

EMBO J. 13, 4278-4290, 1994

A/Title: A novel human protein serine/threonine phosphatase, which possesses four tetrat

A/Reference number: SS2570; MUID:95009929; PMID:7925273

A/Accession: SS2570

A/Molecule type: mRNA

A/Residues: 7-498 <CH2>

A/Cross-references: UNIPROT:P53041; UNIPARC:UPI0000172826; EMBL:S73586; EMBL:X89416

A/Experimental source: teratocarcinoma cell line NTera-2

R/Xu, X.L.; Lagercrantz, U.; Zickert, P.; Bajalica-Lagercrantz, S.; Zetterberg, A.

Biochem. Biophys. Res. Commun. 218, 514-517, 1996

A/Title: Chromosomal localization and 5' sequence of the human protein serine/threonine

A/Reference number: PC4136; MUID:96144708; PMID:8561788

A/Accession: PC4136

A/Molecule type: mRNA

A/Residues: 1-37 <XUX>

A/Cross-references: UNIPARC:UPI000016B012; EMBL:X92121; NID:g1177477; PIDN:CA63089.1; F

A/Experimental source: fetal brain

C/Comment: mRNA encoding this protein was detected in all human tissues tested and was 1

C/Genetics: GDB:PPSC; PPS; PP5

A/Genes: GDB:PPSC; PPS; PP5

A/Cross-references: GDB:136857

A/Map position: 19q13.3-19q13.3

C/Function:

A/Description: catalyzes the hydrolytic dephosphorylation of protein-phosphoserine [val]

A/Note: may play a role in the regulation of RNA synthesis and mitosis

C/Superfamily: protein phosphatase 5; phosphoesterase core homology; phosphoprotein phos

C/Keywords: iron; metalloprotein; nucleus; phosphoric monoester hydrolase; zinc

F/28-61/Domain: tetratricopeptide repeat homology <TT2>

F/62-95/Domain: tetratricopeptide repeat homology <TT3>

F/96-129/Domain: tetratricopeptide repeat homology <TT3>

F/1204-467/Domain: phosphoprotein phosphatase homology <PP>

F/1236-305/Domain: phosphoesterase core homology <PPC>

F/242.244.271/Binding site: iron (Asp, His, Asp) #status predicted

F/271.303.352.426/Binding site: zinc (Asp, Asn, His, His) #status predicted

F/274.304.456/Active site: Asp, His, Tyr #status predicted

F/273.399/Binding site: substrate phosphate (Arg) #status predicted

Query Match 16.5%; Score 567.5; DB 1; Length 498;

Best Local Similarity 34.0%; Pred. No. 4.1e-28;

Matches 144; Conservative 70; Mismatches 142; Indels 67; Gaps 16;

QY 11 FIOKTRRQARREMRRCNMQIFONLEVASBODQALYKFFNDLIKHPQAGRNQYQ 70

Db 95 YIKGYRRAASWALQK-----FR-----AALRD-----YETVVKVPHDKDAKMKQ 137

QY 71 -----GSAHV--SVLDDK-----DVEERFG--DIYAKIKELPIRKNIHL 107

Db 138 ECKKIVKQAFERATAGDEHKASVSDLSISKTIIDETSGPGLDEKVTISFMKE---- 193

QY 108 LIDVFKKRGKGNRLAPKRYVALILREAKSLKOLPNISPVSTAVSQQVTCVGDHKGKLDL 167

Db 194 LMQWYDQK--KLHRCAGVILVQVKEVSLKSTIVETTLKTEKIKTVGSDHGGFYDIL 251

QY 168 VILHKGKGLPSSSNPYFVNDGDRGKGLSLVLLLSLYLAPNAVFLRGNHSDSVMA 227

Db 252 NIFELNGLSEKTNPIFENDFVDRGSGFSEVILTLGFKLLYPDHFLRGNHETDNMQ 311

QY 228 RKGFIRESVSKYPRNKRKLAFIDVYRMLPLGSLVNSVLLVHGFSNST--SLDLIKSI 286

Db 312 IYFEGEYKAKYT--AQMTELFSSEVFWLPLAOCINKKVALIMHGLSVKTYVLDLDIRI 368

QY 287 DRGKYVSLRPLTDEBPLDKTEMQQIFDIMSDDQATMGCVPTLRGAGVWFGPDVTDN 346

Db 369 ERNR-----QPP--DSGPM-----CDLMSDPQNG--RSISKRGVSGQFGFDYTKA 412

QY 347 FLQRRRLSYVRSHCKNGHEFMHNDKTIITFFASNTYALIGSNKAYIRL--NNQLMHP 405

Db 413 FLERNNDLIIRSHVKAQGYVAHGRCTVFSAPNYCDQGNKASYIHLQSGDLRQ 472

QY 406 VQY 408

Db 473 HOY 475

RESULT 7

T45058

phosphoprotein phosphatase (EC 3.1.3.16) Y398B ff [similarity] - Caenorhabditis elegans

N/Alternate names: serine/threonine phosphatase pps homolog

C/Species: Caenorhabditis elegans

C/Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 28-Apr-2003

C/Accession: T45058

R/Wilson, R.; Alnecough, R.; Anderson, K.; Baynes, C.; Berks, M.; Bonfield, J.; Burton, Yasser, A.; Callaghan, M.; Parsons, J.; Percy, C.; Rifken, L.; Koopra, A.; Saunders, D.

Nature 368, 32-38, 1994

A/Authors: Shownkeen, R.; Sims, M.; Smaldon, N.; Smith, A.; Smith, M.; Sonhammer, E.; S

tock, L.; Wilkinson-Sproat, J.; Wohldman, P.

A/Title: 2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.

A/Reference number: 843531; MUID:94150718; PMID:7906398

A/Accession: T45058

A/Status: translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-526 <MIL>

A/Cross-references: UNIPARC:UPI000083436; EMBL:AL132896; NID:96434440; PIDN:CA60937.1;

A/Experimental source: clone Y398B

C/Genetics:

A/Map position: 3

A/Intons: 42/1; 124/3; 184/3; 265/3; 348/3; 385/2; 415/2; 499/3

A/Note: Y398B. ff

C/Superfamily: protein phosphatase 5; phosphoesterase core homology; phosphoprotein phos

C/Keywords: iron; metalloprotein; nucleus; phosphoric monoester hydrolase; zinc

F/29-61/Domain: tetratricopeptide repeat homology <TT1>

F/62-95/Domain: tetratricopeptide repeat homology <TT2>

F/96-129/Domain: tetratricopeptide repeat homology <TT3>

F/203-497/Domain: phosphoprotein phosphatase homology <PP>

F/235-304/Domain: phosphoesterase core homology <PPC>

F/241.243.270/Binding site: iron (Asp, His, Asp) #status predicted

F/270.302.351.456/Binding site: zinc (Asp, Asn, His, His) #status predicted

F/273.303.460/Active site: Asp, His, Tyr #status predicted

F/274.429/Binding site: substrate phosphate (Arg) #status predicted

Query Match 16.2%; Score 558; DB 1; Length 526;

Best Local Similarity 31.6%; Pred. No. 1.7e-27;

Matches 138; Conservative 79; Mismatches 177; Indels 48; Gaps 14;

QY 3 ENAIRPAAL---FIOKTRRQARREMR--RCNMQIFONLEVASBODQALYKFF--FNLDI 56

Db 86 DNAL--ALDPYVVGFGYRATANNALGRFKKALTDYQVAVVYCPNDXARAKFDCSKIV 143

QY 57 KMPQAAKRGQYQGSANVSVDKODLVESFGDIYNAKILPIRKNIHLIDIVFKKR 116

Db 144 RROKFEAISTDHDKGTVAETLDINAMAIIBSYD--GRLDCKITKEFVLQILKIFKQ 201

QY 117 GNRLLPKYVALILREAKSLKOLPNISPVSTAVSQQVTCVGDHKGKLDLVLVHKG 176

Db 202 --KLHKKTAFFKMLEFYVYVSLPTMVEITVPTGKKFTICGVHQGFYDLCKIFETNGY 259

QY 177 SSSNPFVNDGFDVDRGKGLSLVLLLSLYLAPNAVFLRGNHSDSVMA 236

Db 260 SETNPFYLVNDGFDVDRGSGSVETIIFPMIGFKLLYPRHFMHSGNHSDDVMNMGYFGEV 319



[illegible]

RESULT 8  
E84858  
Phosphoenolpyruvate carboxylase (EC 3.1.3.16) At2g42810 [similarity] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 28-Apr-2003  
C:Accession: E84858  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Umayam, L.; Tallon, L.  
euss, D.; Nierman, W.C.; White, O.; Eskin, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,  
Nature 402, 761-766, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: E64858  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-533 <SNO>  
A:Cross-references: UNIPARC:UPI0000044D6A; GB:AB02093; NID:94512673; PTDN:PAD21727.1; C:GeneIDs:  
A:Gene: At2g42810  
A:Map position: 2  
C:Superfamily: protein phosphatase 5; phosphoesterase core homology; phosphoprotein phospho  
C:Keywords: phosphoric monoester hydrolase

```

Query Match      16.2%; Score 557.5; DB 2; Length 533;
Best Local Similarity 39.4%; Pred No. 1.9e-27;
Matches 127; Conservative 55; Mismatches 109; Indels 31; Gaps 10;

QY      AKIE-LPIRKNHIDLLIDVFRKRGGRNRLPHKYVALIILEAAKSILKOLPNISPVSTAVSQO 152
       |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      ARIBEEBVTLDVPVKTMEDPFKNQK--TLHKRYAYIVLQTQIIIALSLVDISTVPHKH 271
       |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY      VTVCGDHLHGKLDLLVLHAKNGLPSSSNPYPFNGDFVDRGKRGLFEVLLLLSLYLAFENA 212
       |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      ITVCGDVAHQGFYDLNLINPELNGLPSEBNPYLFENGGFDVRGSPVEIITLPAFKCMCSS 331
       |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY      VFIAIRGNHEDSVMANARYGFIPIREVESKYPNNHKKLIILAFID---EYVRMLPLGSYLVNSRVLI 269
       |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      IYLAIRGNHSKSMNKITYGFEGEVRSKLSEK-----FYDLPAFAVFCTPLAHVINIGKAFV 385
       |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY      VHGG-FS-DSTSLDIKSIDRGKVYSILRPPLTDSGPLDKTEWQQIPIMMSDPQATGC 327
       |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      VHGGIFSVDSGVKLSIRAIIDR-----PCERP-----BEGLMCLISLSDPPPLGR 430
       |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY      VPNTLRGAAGVWFPGPDVTNPLOQRHRLSYVIRSHCEKPGNHGHEPMHDNKKIITFSASNYAI 387
       |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      GPSK-RGVGLSGFGGVDTKRFLODNNDLLDVSRSHVEKGEGYEVEDGKLIITFSAPNYCDQ 489
       |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY      GSNNGAYTRLN-NQLMPHFVOY 408
       |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      MGNRKAFLRFEADPKKENIVTF 511
       |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

RESULT 9  
T46576  
phosphoprotein phosphatase (EC 3.1.3.16) ptc-1 [similarity] - *Neurospora crassa*  
N:Alternate names: serine/threonine phosphatase PPS homolog

C:Species: Neurospora crassa  
C:Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 28-Apr-2003  
C:Accession: T46576  
R:Yatzkan, B.; Yarden, O.  
Submitted to the EMBL Data Library, February 1997  
A:Description: Ppt-1 a N. crassa novel-type phosphatase.  
A:Accession number: Z23089  
A:Reference: T46576  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-479 <YAT>  
A:Cross-references: UNIPARC:UPI000000D260; EMBL:U89985; PTDN:AA85138.1  
C:Genetics:  
A:Gene: ppt-1  
A:Map position: V  
C:Superfamily: protein phosphatase 5; phosphoesterase core homology; phosphoprotein pho  
C:Keywords: iron; metalloprotein; nucleus; phosphoric monoester hydrolase; zinc  
F:8-1/Domain: tetratricopeptide repeat homology <TT1>  
F:42-75/Domain: tetratricopeptide repeat homology <TT2>  
F:76-109/Domain: tetratricopeptide repeat homology <TT3>  
F:183-448/Domain: phosphoprotein phosphatase homology <PPP>  
F:215-284/Domain: phosphoesterase core homology <PBC>  
F:221,223,250/Binding site: iron (Asp, His, Asp) #status predicted  
F:250,282,331,408/Binding site: zinc (Asp, Asn, His, His) #status predicted  
F:253,283,432/active site: Asp, His, Tyr #status predicted  
F:254,381/Binding site: substrate phosphate (Arg) #status predicted

[illegible]

RESULT 10  
T40391  
Phosphorotransferase phosphatase (EC 3.1.3.16) SPBC3F6.01c [similarity] - fission yeast (Schizosaccharomyces pombe)  
Alternate names: serine/threonine phosphatase pps homolog  
C1Species: Schizosaccharomyces pombe  
C1Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 28-Apr-2003  
C1Accession: T40391  
R1Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Churcher, C.M.  
Submitted to the EMBL Data Library, February 1998

A:Reference number: Z21925  
A:Accession: T40391  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-473 <LYN>  
A:Cross-references: UNIPARC:UPI000006B641; EMBL:AL022019; PIDD:CAA17690.2; GSPDB:GN0006  
A:Experimental source: strain 972h-1, cosmid c3f6  
A:Genetics:  
A:Gene: SPDB:SPBC396.01C  
A:Map position: 2  
C:Superfamily: protein phosphatase 5; phosphoesterase core homology; phosphoprotein phospho  
C:Keywords: iron; metalloprotein; nucleus; phosphoric monoester hydrolase; zinc  
F:35-38/Domain: tetratricopeptide repeat homology <TT1>  
F:39-72/Domain: tetratricopeptide repeat homology <TT2>  
F:73-106/Domain: tetratricopeptide repeat homology <TT3>  
F:119-444/Domain: phosphoprotein phosphatase homology <PPP>  
F:211-280/Domain: phosphoesterase core homology <PEC>  
F:217,219,246/Binding site: iron (Asp, His, Asp) #status predicted  
F:246,278,327,404/Binding site: zinc (Asp, Asn, His, His) #status predicted  
F:249,279,428/Active site: Asp, His, Tyr #status predicted  
F:250,377/Binding site: substrate phosphate (Arg) #status predicted

Query Match 15.1% Score 519; DB 1; Length 473;  
Best Local Similarity 35.2%; Pred. No. 4,3e-25;  
Matches 119; Conservative 66; Mismatches 127; Indels 26; Gaps 8;

QY 73 AHVSLDRKDDLVSEFGDIVNAKTELPFRKHIDLLIDVFAKKGNRLHPKCYVALIREA 132  
DB 139 ANINIED--MIPSPYDQVI---LEKQITKFEVDEMERF---CQKULPKFAYSIARDL 191  
QY 133 AKSLKQLEPNISPVSTAVSQQVTVCGDLHGKLDLLIVLHKNGLPSSSNPVYFNGDPVDRG 192  
DB 192 KELEKTEPSSLIDIPKQGBDETVICDTHQGFYDILLIKLHGPPSPTKVLYFNGDPVDRG 251  
QY 193 KRGLSEVLLLSLVYAFNNAVFLNRGNHEDSVNARVCFIEVSKYPRNKRILATIDE 252  
DB 252 SWSTSEVATLVYAKLLIFDPAVFINRGNHETDDMNKVTFSECSRSKY---NERTFNIFSE 308  
QY 253 VYRWLPFGSVLNSRYLVVHGFG--SDSTSLDLKISIDGKTVSILRPLTDSGEPLDTKEW 310  
DB 309 TFSLLPLFGLSLDSRYLVVHGGLFSPDDNTVTLQGNRIDFS---KKQPGQSG----- 356  
QY 311 QQIFPIMMSDPQATMGCVPTNLRGAGVFPQGVDTNDFLQRRRLSYVIRSHCECKNGHEFM 370  
DB 357 -LWMEMLMTDPPAPGRGRPSK-RGVGLQFGPDVSKRFEANGLXAVIRSHSEVRDQGYEVE 414  
QY 371 HDNKLITTFPSASNYAIAISNGKGYVRLNOLMHPHYQY 408  
DB 415 HDGYCTIVFSAPNYCDSTGNLDAVIVKXEDMELDHPQ 452

RESULT 11  
S52571  
phosphoprotein phosphatase (EC 3.1.3.16) PPT1 - yeast (Saccharomyces cerevisiae)  
N/Alternate names: protein G6347; protein YG8123c; serine/threonine phosphatase PPS homoc  
C/Species: Saccharomyces cerevisiae  
C/Date: 03-Aug-2001 #sequence revision 03-Aug-2001 #text\_change 09-Jul-2004  
C/Accession: S52571; S55981; S64432; S64697  
R/Chen, M.X.; McPartlin, A.E.; Brown, L.; Chen, Y.H.; Barker, H.M.; Cohen, P.T.W.  
EMBO J. 13, 4278-4290, 1994  
A/Title: A novel human protein serine/threonine phosphatase, which possesses four tetra  
A/Reference number: S52570; MUID:95009929; PMID:7925273  
A/Accession: S52571  
A:Molecule type: mRNA  
A:Residues: 1-513 <CHR>  
A:Cross-references: UNIPROT:P50403; UNIPARC:UPI0000053298  
R/ivan Dyck, L.; Goffeau, A.  
submitted to the EMBL Data Library, December 1994  
e new ORFs, remnants of Ty and three tRNA genes.  
A:Reference number: S55976  
A/Accession: S55981  
A:Molecule type: DNA

A:Residues: 1-513 <VAM>  
A:Cross-references: UNIPARC:UPI0000053296; EMBL:X83099; NID:G642340; PID:G642346  
R:Van Dyck, L.; Skala, J.; de Weyfossé, P.; Purnelle, B.; Talla, E.; Nawrocki, A.; Del  
submitted to the Protein Sequence Database, May 1996  
A:Reference number: S64428  
A:Accession: S64432  
A:Molecule type: DNA  
A:Residues: 1-513 <VAM>  
A:Cross-references: UNIPARC:UPI0000053296; EMBL:Z72908; NID:G133200; PID:G243462; PID:G  
A:Experimental source: strain S288C  
R:Chen, M.X.; McPartlin, A.E.; Brown, L.; Chen, Y.H.; Baker, H.M.; Cohen, P.T.W.  
submitted to the EMBL Data Library, July 1995  
A:Description: A novel human protein serine/threonine phosphatase, which possesses four  
A:Reference number: S64697  
A:Accession: S64697  
A:Molecule type: DNA  
A:Residues: 1-380, 'HLE', 369, 'AP', 392-393, 'NFGRTTKKPGWTL', 409, 'STV', <CHW>  
A:Cross-references: UNIPARC:UPI0000168BI; EMBL:X89417; NID:G897805; PID:G897806  
C:Genetics:  
A:Gene: SGD:PPY1  
A:Cross-references: SGD:S0003355; MIPS:YGR123C  
A:Map position: 7R  
C:Superfamily: protein phosphatase 5; phosphoesterase core homology; phosphoprotein phos  
C:Keywords: Iron; metalloprotein; nucleus; phosphoric monoester hydrolase; zinc  
F:12-45/Domain: tetratricopeptide repeat homology <TT1>  
F:46-79/Domain: tetratricopeptide repeat homology <TT2>  
F:80-113/Domain: tetratricopeptide repeat homology <TT3>  
F:209-475/Domain: phosphoprotein phosphatase homology <PPP>  
F:243-312/Domain: phosphoesterase core homology <PES>  
F:243,251,778/Binding site: Iron (Asp, His, Asp) #status predicted  
F:278,310,359,434/Binding site: zinc (Asp, Asn, His, His) #status predicted  
F:281,311,458/Active site: Asp, His, Tyr #status predicted  
F:282,407/Binding site: substrate phosphate (Arg) #status predicted

Query Match 13.9%; Score 479.5; DB 1; Length 513;  
Best Local Similarity 32.3%; Pred. No. 1,46-22;  
Matches 130; Conservative 68; Mismatches 154; Indels 51; Gaps 13;

2 DENAIRAIRITOKYRRHQAARK--QRRCNMQIFQNTLYASRQDAELYKFFNDLIK 58  
113 DPATIKALLCDRIPIRERFRKAIQGAENAKISLCOTLNLSPFADNADLANV----- 165  
59 MPQAQKNGYQGSAAVSLDDKDLVEFEGDIVNAKISLPKRNHIDLLIDVFRKKGN 118  
166 ---EGPKLSEF-----QLYDDKN-----AFKAKIKIKMGSEFISKVNDLFLK--GK 207  
119 RLRHKYVALLLRBAKSLKQLPNISPV--STAVSQOYVTCGLHGLKLDLLVTLHKGGLP 176  
208 YLPCKYVAAIISHADTLFRQEPSEVLENNSTPDVAKISVCGDTHGQFYDVLNFRFKGV 267  
177 SSSNPYFQNDPVDGRKGLJEVLLLSLYLAFPNVAFVLRGNHDSVNNARAGFYREVE 236  
268 GPKHTYLPNDPVDGRGWSCEVALLFYCLKILIPNFLLRGNHDSNNMKIYGFEDCK 337  
237 SKYPRNKRRIAFIDEVYRMLPGSVLNSKVLIVHGQF--SD--STSIDLILKISGRKVS1 294  
328 YKY---SGRFNNPAGSFSFLPLATILNNDYLVHGGFLPDPBATILSDFNKIDR-----F 379  
295 LRPLDTGEBLDTKEMQCIDIDIMSDPOATMGCVPTNLRGAGVWFGPDVTNLTQRRHS 354  
380 AQP-RLG-----AFMELLMADPOEANGKPSQ-RLIGHAFGPDIDRLRNKLR 428  
355 VYIRSHCKDNGHFMEDNKIITFSASNTYVYAGSNKGAIVRL 397  
429 KIFRSHLRMGVQVFEQKGLMTVFSAPNYCDGQNGDVIVH 471

RESULT 12  
T51611  
phosphoprotein phosphatase (EC 3.1.3.16) P7 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 09-Jul-2004  
C:Accession: T51611

R.Andreeva, A.V.; Evans, D.B.; Hawes, C.R.; Bennett, N.; Kutuzov, M.A.

Biochem. Mol. Biol. Int. 44, 703-715, 1998

A/Title: PP7, a plant phosphatase representing a novel evolutionary branch of eukaryotic

A/Reference number: Z25411

A/Accession: T51611

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1413 <AND>

A/Cross-references: UNIPROT:Q49346; UNIPARC:UP100000ACCB3; EMBL:AJ000057; P1DN:CA03886;

A/Experimental source: cultivar Columbia; whole plants; 3 weeks old; grown in liquid med

C/Function:

A/Description: catalyzes the hydrolytic dephosphorylation of protein-phosphoserine and p

C/Keywords: phosphoric monoester hydrolase; serine/threonine-specific phosphatase

Query Match 12.9%; Score 444; DB 2; Length 413;  
Best Local Similarity 35.6%; Pred. No. 1.8e-20;  
Matches 117; Conservative 38; Mismatches 98; Indels 76; Gaps 10;

138 QLEPISPVSTAVS-----QQVTGGLHGKLDLLVYLHN 173

40 QLESLPVPVPSLVLTAKILHKERNCHIDLDLSVSNVVGVGDHGLDILLFLKOT 99

174 GLPSSNPYVNGDFVDRGKGLVLLLSLYLAPNAVFLRGNHDSVMNARYGPIR 233

100 GPFQCRVCYFNGDYDRGMGLFTPLVLSKVLMPDRVYLARKHESKCYTSMTGFER 159

234 EYESKTPRNHRIIAFIDEVYRWLPGLSVNSVLIHGGFSDSTSLDIKSIDRK-- 290

160 EVLTGKDGKHYRKLGLCFEGSLPLASITSGRYAHAGLFPSPVLP--KRTTRGKOR 217

291 YSLAPPLDGR-----LDK-----TEMQO-----PIMSDPQATMGCP 329

218 RVLLP-----PSSMKLGTLDLMQARRSVLDPMWGSNLLPGVLMSPSITGGLSP 272

330 NTLRAGVWGPVDTNPLQRHRLSYVIRSH-----CKNGHGFWD--NK 374

273 NERGIGLMGPRCTEDPLKYLKILIRSHBGPDAERKRTGSGMDNGTIDHNVESGK 332

375 IITIFASNYAIGS-----NKGAYIRL 397

333 LITIFAPDPQVQATVEERYKNKGAVIIL 361

RESULT 13  
T09995  
phosphoprotein phosphatase (EC 3.1.3.16) 1a catalytic chain - Madagascar periwinkle  
N/Alternate names: PPIA protein  
C/Species: Catharanthus roseus (Madagascar periwinkle)  
C/Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
C/Accession: T09995  
R/Camilleux, M.; Chenieux, J.C.; Rideau, M.; Hamdi, S.  
Plant Physiol. 118, 1533, 1998  
A/Title: cDNA isolation of two serine/threonine phosphatases (PPI1 and PPI2) from Cathar  
A/Reference number: Z16912; MIMD:99139832; PMID:9988639  
A/Accession: T09995  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-326 <CAN>  
A/Cross-references: UNIPROT:Q81955; UNIPARC:UP100000A8113; EMBL:AJ007332  
A/Experimental source: cell line C20D  
C/Genetics:

A/Gene: PPIA  
C/Function:  
A/Description: catalyzes hydrolysis of peptidyl-phosphoserine or -phosphothreonine to re  
C/Superfamily: serine/threonine protein phosphatase; phosphoesterase core homology; phos  
C/Keywords: phosphoric monoester hydrolase; serine/threonine-specific phosphatase  
P/26-285/Domain: phosphoprotein phosphatase homology <PPP>  
P/54-122/Domain: phosphoesterase core homology <PPC>

Query Match 12.8%; Score 441; DB 2; Length 326;  
Best Local Similarity 32.0%; Pred. No. 2e-20;  
Matches 103; Conservative 64; Mismatches 125; Indels 30; Gaps 7;

105 IDLLIDVPRKRGK-----RLHPKVALLREAAKSLKQLPNISPVSTAVSQVTCGL 159

6 LGIINRLIIVRGPRKQVQLSESRIRHLCLKSKIRFMQPNLIEAPK-----ICGDI 61

160 HGLDLDLVLVHKNGLPSSSNPYVNGDFVDRGKGLVLLLSLYLAPNAVFLRGN 219

62 HQGYDLRLRPYGGVGLPSPN--YLPGLGVYDVGKQSLFTICLLAKIKYIPKIFLRKN 120

220 HEDSVNNAARYGPIREVSRYPRNHRRIIAFIDEVYRWLPGLSVLNSVLIHGGFS-DST 278

121 HECASVNRYYGYGECKRPF--NVALMKVFTDCEFCFLPVALIBKILCMHGLSPDIL 177

279 SLDLIKSIDRGKYSILRPPLDGBRLDTEWQOIFDIMWSPQATMGCVPTLRAGTW 338

178 SLDQIRNLPR-----PTDVEGTGLCDLWSPDSKQVKGMDNRGVSTY 222

339 PGPDVTNPLQRHRLSYVIRSHCKPNGHFMNDKTIITIFASNYAIGSNKAYIRLN 398

223 PGPDVKKEFLQKHLDLIGRAHQVVEDGYEFPADRLQVITIFAPNYCGSPDNAGMMSVD 282

399 NQMPHFVQYISAAGTKRLSP 420

283 ETLMCSP-QILNPAEKAKFGF 303

RESULT 14

S52371  
phosphoprotein phosphatase (EC 3.1.3.16) PPI - kidney bean

C/Species: Phaseolus vulgaris (kidney bean)  
C/Date: 08-May-1995 #sequence\_revision 03-Nov-1995 #text\_change 09-Jul-2004  
C/Accession: S60053; S52371  
R/Zimmerlin, A.; Ujpe, S.C.; Bolwell, G.P.  
Plant Mol. Biol. 28, 363-368, 1995

A/Title: Molecular cloning of the cDNA encoding a stress-inducible protein phosphatase

A/Reference number: S60053; MIMD:95359396; PMID:7632908

A/Accession: S60053

A/Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tr

A/Molecule type: mRNA

A/Residues: 1-314 <Z12>

A/Cross-references: UNIPROT:P48490; UNIPARC:UP10000131FC0; GB:248221; NID:9667048; P1DN:

C/Superfamily: serine/threonine protein phosphatase; phosphoesterase core homology; phos

C/Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-speci

P/24-283/Domain: phosphoprotein phosphatase homology <PPC>

P/52-120/Domain: phosphoesterase core homology <PPC>

P/58-60/Binding site: iron (Asp, His, Asp) #status predicted

P/86-118,167,242/Binding site: zinc (Asp, Asn, His, His) #status predicted

P/89,119,266/Active site: Asp, His, Tyr #status predicted

P/90,215/Binding site: substrate phosphate (Arg) #status predicted

Query Match 12.7%; Score 436.5; DB 2; Length 314;

Best Local Similarity 31.5%; Pred. No. 3.7e-20;

Matches 105; Conservative 73; Mismatches 116; Indels 39; Gaps 10;

107 LLDIVPRK-----KRGRLHPKVALL--LREAAKSLKQLPNISPVSTAVSQVTCGL 159

5 VLDDVIRLLRBSGG-----KQVQLSESRIRQLCVARQIPUSQPLLELRAPMRVCGDI 59

160 HGLDLDLVLVHKNGLPSSSNPYVNGDFVDRGKGLVLLLSLYLAPNAVFLRGN 219

60 HQGYDLRLRPYGGVGLPSPN--YLPGLGVYDVGKQSLFTICLLAKIKYIPKIFLRKN 118

220 HEDSVNNAARYGPIREVSRYPRNHRRIIAFIDEVYRWLPGLSVLNSVLIHGGFS-DST 278

119 HEEAKINRIYGYDECKRPF--NVALMKVFTDCEFCFLPVALIBKILCMHGLSPDIL 175

279 SLDLIKSIDRGKYSILRPPLDGBRLDTEWQOIFDIMWSPQATMGCVPTLRAGTW 338

176 SLDQIRNLPR-----PTDVEGTGLCDLWSPDSISGMASSDRGVSTCT 220

339 PGPDVTNPLQRHRLSYVIRSHCKPNGHFMNDKTIITIFASNYAIGSNKAYIRLN 398

221 PGPDVKKEFLQKHLDLIGRAHQVVEDGYEFPADRLQVITIFAPNYCGSPDNAGMMSVD 280







R.M.I.D. A.  
submitted to the EMBL Data Library, July 1996  
A.Reference number: Z19402  
A.Accession: T21288  
A.Status: translated from GB/EMBL/DBJ  
A.Molecule type: DNA  
A.Residues: 1-401 <MIL>  
A.Cross-references: UNIPROT:Q27495; UNIPARC:UPI000007C7BF; EMBL:Z77659; PIDN:CAB01164.1;  
A.Experimental source: clone F23B12  
C.Genetics:  
A.Gene: CESP:F23B12.1  
A.Map position: 5  
A.Introns: 71/3; 121/3; 153/2; 251/2; 377/2  
C.Superfamily: serine/threonine protein phosphatase; phosphoesterase core homology; phos  
C.Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-speci  
F.197-356/Domin: phosphoprotein phosphatase homology <PP>  
F.125-193/Domin: phosphoesterase core homology <PEC>  
F.131.133.159/Binding site: iron (Asp, His, Asp) #status predicted  
F.159.191.240.315/Binding site: zinc (Asp, Asn, His, His) #status predicted  
F.162.192.339/Active site: Asp, His, Tyr #status predicted  
F.163.288/Binding site: substrate phosphate (Arg) #status predicted

Query Match 12.3%; Score 425; DB 1; Length 401;  
Beet Local Similarity 29.3%; Pred. No. 2.7e-19;  
Matches 113; Conservative 68; Mismatches 137; Indels 68; Gaps 12;

OY 50 KEFNDD-----IKHMPQAARKKQOGSAHVSYLDDEDIVBEFGDIVNAKIBLPYRN 103  
||| ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 13 KFGDVPHVDLFPRPPLSINQNSIY---SSVSL-SKSSMEYR-----ESSAISIRN 63  
||| ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
OY 104 HIDLIDVFRKKRGNRLLHP-----KYVALIREANSLKQLPNIS 143  
||| ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 64 AAG--NDVOKNKQNMHLSTYERLKMWSPGRCOQLFVENELIELCYRAREGEFWK-----N 116  
||| ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
OY 144 PVTANSQQVTGCGDLHGKLDLLVLHKNLBPSSSNYYVNGDPDVGKGLFTYLILL 203  
||| ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 117 KYKLDEAPFKITGGDIHGQEDLMALFEINLWNP-BEHKLYLPGDVDGPFSEIVYTLLF 175  
||| ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
OY 204 SYLAEPNAVFLRGNHSDVMNARYGFIREVESKYPNNHRILLAFIDE---VYRWPLTG 260  
||| ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 176 TEOILMPDKVFLRGNHESRPVMQGYFLECKKRYT-----VALYDAFOALPCMLPC 229  
||| ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
OY 261 SVLNSRVLLVHGFS-DSTSLDLKSIDRGKVSLRLPPLTDGBEPLDTIEWOOIPIDWMS 319  
||| ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 230 AVASKKIICMGHSISBDLIDLTOLKEIDR-----PFDLPDIGVISDLTMA 274  
||| ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
OY 320 DPOATMGCPVNTLRGAWGFGPDVTNPLQRHLSTYRSHECKNGHEPMHNKITTF 379  
||| ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 275 DDEKEFGYADSPRGGRSGFPNNAVKFFLOQHNLIDLVRHQVVMDGYEFPADROLVTVF 334  
||| ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
OY 380 SASNYVAIGSKNGAYIRLNQMLMPHF 405  
||| ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 335 SAPSYCGQDFDNAAAAMNVMDKLLCTF 360  
||| ||| : | : | : | : | : | : | : | : | : | : | : | : | :

RESULT 23  
S31086  
phosphoprotein phosphatase (EC 3.1.3.16) 1 catalytic chain (clone TOPP2) - Arabidopsis th  
C.Species: Arabidopsis thaliana (mouse-ear cress)  
C.Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004  
C.Accession: S31086  
R:Smith, R.D.; Walker, J.C.  
Plant Mol. Biol. 21, 307-316, 1993  
A>Title: Expression of multiple type 1 phosphoprotein phosphatases in Arabidopsis thalia  
A.Reference number: S31085; MUID:J3144705; PMID:7678768  
A.Molecule type: mRNA  
A.Residues: 1-312 <SMI>  
A.Cross-references: UNIPROT:P48482; UNIPARC:UPI000008D80; EMBL:M93409; NID:g166796; PFI  
C.Superfamily: serine/threonine protein phosphatase; phosphoesterase core homology; phos  
C.Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-speci  
F.136-295/Domin: phosphoprotein phosphatase homology <PP>  
F.64-132/Domin: phosphoesterase core homology <PEC>

Query Match 12.3%; Score 424.5; DB 2; Length 312;  
Beet Local Similarity 30.5%; Pred. No. 2.1e-19;  
Matches 109; Conservative 63; Mismatches 128; Indels 57; Gaps 10;

68 OYQSAHVAVLDDKDDLVEEFGDILVNAKILPIRKNHIDILIDVFRKKRG-----NRLL 121  
4 OQGSMDPALDD-----ILRR-----LIDYRNPFGTKQALNISE 40

122 PRYVALIREAAKSLKQLPNISPVSTAVSQOVTVCGLHGKLDLVLVLRKNGLPSSNP 181  
41 IRLQICIVSREI---FLQGFNLLELBAPIK-----ICGDIHQGYSLDLRLFEVGGPEPTAN- 92

182 YTFNGDPVDNRGKRGLEVLILLLSLYLAFPNAVFLNRGNHEDSVNNARVGFIREVSKYPR 241  
93 YLFLDGYVDNRGKQSLLETCILLAYKIKYPENFPLNRNHCASINRIIGFYDECKRRF-- 150

242 NHRILAFIDEVYRWMLPLGSVLSRVLIYHGFS-DSTSLDLSIDRGKVSILRPPLT 300  
151 -SVRLMKVFTDSNCLPVAIVIDKILCMHGGLSPDLINVEQIKNIK----- 197

301 DEEPLDKTEWOOIFDIWMSDPCQATMGCVPTLKGAGVWFGPDYVDNPLQRRHLSYVIRSH 360  
198 ---PFDVPSGLCDLMDSPDKVDKMGKMDNVGSYTFPGDXKAFILKNDMDLICAH 254

361 ECKPENGHEMMDNKITITTSASNYVAIGSKKAVIRLNNLMHPVOYISAAQTKR 417  
255 QVVEGVEEFPADRLQVLTFISAPNYCGEFDNAGAMSVDESIMCSF-QILKPADRKPR 310

RESULT 24  
T09550

phosphoprotein phosphatase (BC 3.1.3.16) 1, catalytic epsilon chain - alfalfa  
C:Species: Medicago sativa (alfalfa)  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
C:Accession: T09550

R:Vissel, B.; Coorides Toth, E.; Kovacs, I.; Magyar, Z.; Horvath, G.; Bagossi, P.; Gergely, Arch. Biochem. Biophys. 360, 206-214, 1998  
A:Title: Protein phosphatase 1 catalytic subunit isoforms from alfalfa: Biochemical char  
A:Reference number: Z16730; MUID:99068922; PMID:9851832  
A:Accession: T09550

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-333 <VITS>  
A:Cross-references: UNIPROT:O6S847; UNIPARC:UP100000AABFC; EMBL:AJ002488; NID:g3176077;  
A:Experimental source: subspecies sativa; strain RA3  
C:Genetics:

A:Gene: PPI epsilon  
A:Description: catalyzes hydrolysis of peptidyl-phosphoserine or -phosphothreonine to r  
C:Keyword: serine/chreonine protein phosphatase; phosphoesterase core homology; pho  
C:Keywords: phosphoric monooester hydrolase; serine/chreonine-specific phosphatase  
F:42-301/Domain: phosphoprotein phosphatase homology <PPP>  
F:70-138/Domain: phosphoesterase core homology <PPEC>

Query Match 12.2%; Score 421; DB 2; Length 323;  
Beet Local Similarity 32.5%; Pred. No. 3.6e-19;  
Matches 106; Conservative 58; Mismatches 132; Indels 30; Gaps 7;

107 LLIIDVFRKKGRNRLHAKYVALIIREAAKSLKQLPNIS-----PVSIVSQOVTVCGLD 159  
20 VLDIDIRRLTEVRLSPRKQVOLSEK--EIKOLCSASRDIFLQOPNLELEAPITICDI 77

160 HGKLDLILVLRKNGLPSSNPVYVFGDPVDNRGKRGLEVLILLLSLYLAFPNAVFLNRGN 219  
78 HCGYSDLLRLFEVGGIIPQGN-YLFLGDIYDRGKQSLLETCILLAYKIKYKIEENFLILGN 136

220 HEDSVNNARVGFIREVSKYPRNHRKILAFIDEVYRWMLPLGSVLSRVLIYHGFS-DST 278



Db 137 HECASINRIYGFYDECKRRF---NVALMKAFIESFNCFLVAAALIDKILCMHGSLPDLT 193  
Qy 279 SLDLISIDRGKVVSLRPPLDGEPDLDTMWOQIFDIMSPOATMGCVPTLTGAGTW 338  
Db 194 NIDQJINLPR-----VPPIDPTGLRDQFMSDKGVKCMGMANDGVSYT 238  
Qy 339 FGPDVTDNFQIRHLSYVIRSHCEKNGHEFPMHDKITITFASNYAIAIGSKGAYIRLN 398  
Db 239 FGPDKAAEFITRNDLDICRAHQVAVDGTFFPADROLVITFSAFNTCGEFNDAGAMMSVD 298  
Qy 399 NQIMPHFVOYISAASQTKSLSPKQRM 424  
Db 299 ENLMCSF-QILKPAKPKTKFVMSNKM 323

RESULT 25  
T03548  
phosphoprotein phosphatase (EC 3.1.3.16) 1 catalytic chain delta - alfalfa  
C/Species: Medicago sativa (alfalfa)  
C/Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
C/Accession: T03548  
R/Viabel, E.; Csootdae Toch, E.; Kovacs, I.; Magyar, Z.; Horvath, G.; Bagossi, P.; Gergely Arch. Biochem. Biophys. 360, 206-214, 1998  
A/Title: Protein phosphatase 1 catalytic subunit isoforms from alfalfa: Biochemical char  
A/Reference number: Z16730; MUID:99068922; PMID:9851832  
A/Accession: T03548  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-326 <VIS>  
A/Cross-references: UNIPROT:O65846; UNIPARC:UPI00000AAB92; EMBL:AJ002487; NID:G3176075;  
A/Experimental source: subspecies sativa, strain RA3  
C/Genetics:  
A/Gene: PPI delta  
C/Function:  
A/Description: catalyzes hydrolysis of peptidyl-1-phosphoserine or -phosphothreonine to re  
C/Superfamily: serine/threonine protein phosphatase; phosphoserine core homology; phos  
C/Keywords: phosphoric monoester hydrolase; serine/threonine-specific phosphatase  
F/26-285/Domain: phosphoprotein phosphatase homology <PPP>  
F/54-122/Domain: phosphoserine core homology <PESC>

Query Match 12.2%; Score 421; DB 2; Length 326;  
Best Local Similarity 30.9%; Pred. No. 3.6e-19;  
Matches 104; Conservative 67; Mismatches 128; Indels 36; Gaps 9;

Qy 105 IDLLIVPFRKGNRLHPKVALILRBAKSLKQLENIS-----PVSITVSGQVTVCG 157  
Db 10 INRLLELV-RGRPEKQVQ-----LSBA--EIKQLCLVSKDIFPMQPNILKEAPIKICG 59  
Qy 158 DLHGKLDLLVVLHKNGLPSSSNPVYFNGDPVDVGRGKEVLLILSLVLAIPNANVFLNR 217  
Db 60 DIHGQYSDLRLEFYGGFPPRSH-YLFLGDYVDKGQSLSTICLLAYIKIKIPKQFFLR 118  
Qy 218 GNHEDSVNARVYGFIREVESKYPNRHRIILAFIDVYRWPLGSLVNSRVLIVHGFS-D 276  
Db 119 GNHGASINRIYGFYDECKRRY---NVKLMQMTDFCNCLPVAALIDKILCMHGSLSP 175  
Qy 277 STSLDLISIDRGKVVSLRPPLDGEPDLDTMWOQIFDIMSPOATMGCVPTLTGAG 336  
Db 176 LHDILNLIKMLR-----PCFVPSGGLCDLMSDPSDVRGMGSRSGVS 220  
Qy 337 VWFPGVTDNFQIRHLSYVIRSHCEKNGHEFPMHDKITITFASANYAIAIGSKGAYIR 396  
Db 221 YTFGADRVKVEFLCKHLDLDCRAHQVAVDGTFFPADROLVITFSAFNTCGEFNDAGAMMT 280  
Qy 397 LNNQIMPHFVOYISAASQTKSLSPKQRMGIVSSALK 433  
Db 281 VNESLVCSF-QILKPDKPKKFSFGSTTYKSSSPFK 316

RESULT 26  
T03304  
probable phosphoprotein phosphatase (EC 3.1.3.16) 1 catalytic chain - rice  
C/Species: Oryza sativa (rice)

C/Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004  
C/Accession: T03304  
R/Wang, B.; Chang, M.; Chen, X.; Wu, R.  
submitted to the EMBL Data Library, July 1995  
A/Reference number: Z14900  
A/Accession: T03304  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-317 <MAN>  
A/Cross-references: UNIPROT:P48489; UNIPARC:UPI0000131BF; EMBL:U31773; NID:G951335; PID  
A/Experimental source: seed  
C/Superfamily: serine/threonine protein phosphatase; phosphoserine core homology; phos  
C/Keywords: phosphoric monoester hydrolase  
F/36-295/Domain: phosphoprotein phosphatase homology <PPP>  
F/64-132/Domain: phosphoserine core homology <PESC>

Query Match 12.2%; Score 420.5; DB 2; Length 317;  
Best Local Similarity 31.6%; Pred. No. 3.8e-19;  
Matches 103; Conservative 61; Mismatches 123; Indels 39; Gaps 10;

Qy 108 LIDVFRKRGRLHPKVALILRBAKSLKQLENIS-----PVSITVSGQVTVCGDLH 160  
Db 22 LLEVTRAPRGQVH-----LSBS--EIKQLCVSRREIFLSQPLILRLKAPIKICGDIH 72  
Qy 161 GKLDDLLVVLHKNGLPSSSNPVYFNGDPVDVGRGKEVLLILSLVLAIPNANVFLNR 220  
Db 73 QQYSDLLLFYGGFPPRSH-YLFLGDYVDKGQSLSTICLLAYIKIKIPKQFFLR 131  
Qy 221 EDSVNNARVYGFIREVESKYPNRHRIILAFIDVYRWPLGSLVNSRVLIVHGFS-DSTS 279  
Db 132 HECASINRIYGFYDECKRRF---NVALMKAFIESFNCFLVAAALIDKILCMHGSLPDLT 188  
Qy 280 LDIKSIDRGKVVSLRPPLDGEPDLDTMWOQIFDIMSPOATMGCVPTLTGAGTW 338  
Db 189 LDEIRKLAR-----PTDVPDTGLCDYFMSDVGKQVQGMGMND-RGVSYT 232  
Qy 339 FGPDVTDNFQIRHLSYVIRSHCEKNGHEFPMHDKITITFASNYAIAIGSKGAYIRLN 398  
Db 233 FQADVSEFLEHGDLDLCRAHQVAVDGTFFPADROLVITFSAFNTCGEFNDAGAMMSVD 292  
Qy 399 NQIMPHFVOYISAASQTKSLSPKQRM 424  
Db 293 ENLMCSF-QILKPAKPKTKFVMSNKM 317

RESULT 27  
T21322  
phosphoprotein phosphatase (EC 3.1.3.16) 1 P25B3.4 [similarity] - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 05-Oct-2004  
C/Accession: T21322  
R/Gardner, A.  
submitted to the EMBL Data Library, April 1996  
A/Reference number: Z19406  
A/Accession: T21322  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-291 <WIL>  
A/Cross-references: UNIPROT:Q27496; UNIPARC:UPI000007D589; EMBL:Z70752; PIDN:CAA94756.1;  
A/Experimental source: clone F25B3  
C/Genetics:  
A/Gene: CESP:P25B3.4  
A/Map position: 5  
A/Introm: 20/2; 157/3; 273/3  
C/Superfamily: serine/threonine protein phosphatase; phosphoserine core homology; phos  
C/Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-speci  
F/7-268/Domain: phosphoprotein phosphatase homology <PPP>  
F/35-103/Domain: phosphoserine core homology <PESC>  
F/41.43.69/Binding site: iron (Asp, His, Asp) #status predicted  
F/69.101.151.227/Binding site: zinc (Asp, Asn, His, His) #status predicted  
F/72.102.251/Active site: Asp, His, Tyr #status predicted  
F/73.200/Binding site: substrate phosphate (Arg) #status predicted



Query Match 12.2%; Score 420; DB 1; Length 291;  
Beet Local Similarity 35.2%; Pred. No. 3.6e-19;  
Matches 94; Conservative 48; Mismatches 107; Indels 18; Gaps 6;

Dy 141 NISPVSTAVSQQVTVCGDLSHGKLDLVLVHKNGLPSSSNPFVPFGDFDRGRGLEVLL 200  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db NKQPMLEIVNSPINICGDHHCQFSDLRFLPKNGPRHAN YLFGLDYDRKKCLJETII 82

Dy 201 LILSLYLAEPNAVPLNKGNHEDSVNARVGFIREYSKYPRNNKRILLAFIDVYRWLPILG 260  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db LLFAVAYVFENHFEMLRGNHCSLINRQGYEVECCORY--NKPSWHSFGGVFSVMPILT 140

Dy 261 SVLVNSVLYVHGERSDSITDLIKSIDRGKVSIILRPVLTDGEPRDKTMOQLFDIMWSD 320  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db ALVGQRILLCHMGVS-----TKLNVSQRLAKRFPEMDBPPTLA-----IDLMSD 187

Dy 321 P-QATNGCVNTIRGAGVWFPGPDVTNDPLQRRRLSYVRSHCKNGEHFMDNKIITTF 379  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db PTNFQGMGNPNS-RGVSYVGSDDLRLKLDDIQLDVLVRAHQVDGIEFFANRRLLVTF 246

Dy 380 SASNYTAIGSNKGAAYIRLNNQMPPHV 406  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db SAPFYCGFDMAAAMVVYNKVLCSFV 273

RESULT 28

S31088  
Nonphosphoprotein phosphatase (EC 3.1.3.16) 1 catalytic chain (clone TOPP4) - Arabidopsis thaliana  
Alternative names: hypotheical protein Atg39840; protein T517.14  
C Species: Arabidopsis thaliana (mouse-ear cress)  
C Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004  
C Accession: S31088; T01014; B84822  
R Smith, R.D.; Walker, J.C.  
Plant Mol. Biol. 21, 307-316, 1993  
A Title: Expression of multiple type 1 phosphoprotein phosphatases in Arabidopsis thaliana  
A Reference number: S31085; MUID:53144705; PMID:7678768  
A Accession: S31088  
A Molecule type: mRNA  
A Residues: 1-321 <RNA>  
A Cross-references: UNIPROT:P48484; UNIPARC:UPI0000131FB4; EMBL:M93411; NID:9166800; PubMed:1106231/Binding site: substrate phosphate [Arg] #status predicted

Rounslay, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul submitted to the EMBL data library, November 1997  
A Description: Arabidopsis thaliana chromosome II BAC T517 genomic sequence.  
A Reference number: Z14162  
A Accession: T01014  
A Status: translated from GB/EMBL/DDBJ  
A Molecule type: DNA  
A Residues: 1-321 <RNA>  
A Cross-references: UNIPARC:UPI0000131FB4; EMBL:AC003000; NID:92642152; PIDN:AAB87136.1; PubMed:1106231/Binding site: substrate phosphate [Arg] #status predicted

A Experimental source: cultivar Columbia  
R Lin, X.; Kaul, S.; Rounslay, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Umayam, L.; Tallon, L.; Neus, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, A. Nature 402, 761-768, 1999  
A Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A Reference number: AB4420; MUID:20083487; PMID:10617197  
A Accession: B84822  
A Status: preliminary  
A Molecule type: DNA  
A Residues: 1-321 <RNA>  
A Cross-references: UNIPARC:UPI0000131FB4; GB:AB002093; NID:92642169; PIDN:AAB87136.1; GeneAccess:  
A Gene: Atg39840; T517.14  
A Map position: 2  
A Intron: 73/1; 259/3  
C Superfamily: serine/threonine protein phosphatase; phosphoesterase core homology; phos C Keywords: iron, metalloprotein; phosphoric monoester hydrolase; serine/threonine-spec F:40-299/Domain: phosphoprotein phosphatase homology <PPP>  
F:68-136/Domain: phosphoesterase core homology <PC>  
F:74.76.102/Binding site: iron (Asp, His, Asp) #status predicted  
F:102.134.183.258/Binding site: zinc (Asp, Asn, His, His) #status predicted  
F:103.135.282/Active site: Asp, His, Tyr #status predicted  
F:106.231/Binding site: substrate phosphate [Arg] #status predicted

```

Query Match      12.2%; Score 419.5; DB 2; Length 321;
Best Local Similarity 32.3%; Pred. No. 4.4e-19;
Matches 104; Conservative 56; Mismatches 123; Indels 37; Gaps 8;

Oy 107 LIDVYRKRRKGNLHPKYVALILREAAKSLLQQL-----NISEPVSTAVSQOVTV 155
Db 18 VLDIDIIRLLTRVLRAVGQVQLSEA--EIKQLCTTADIFLQGNLELAPIK----I 71
Oy 156 CGDLHGKLDLVLVHLKNGLPSSSNPYFVNGDFVDGRKGGLVLLLISLYLAFPNAVPL 215
Db 72 CGDHIHQYSIDLTLRFXYGGFFPPSAN-YLFLDDYVDRGQSLETITCLLAAYKITKYPGNAFPL 130
Oy 216 NRGNHSDSVNNARYGFIREVESKYPRNMKRILLAFIDEYYRMPLGSVINSRVLIVAGPS 275
Db 131 LGNHNCASINRIYGFEDCKRRP---NVRYMKVPTDFCNCLPVALLIDDKTLCMHGGLS 187
Oy 276 -DTSIDLILKSIDRGKVSYILRPPLTDGEPLDKTEWOOIPDIMSDPGATMGCVPNTLRG 334
Db 188 PDLDDLDEIRNLPE-----RPTMIPTDGL-----LCOLLMSDPGDVGVGMKNNDG 232
Oy 335 AGYWFEPDVTDNFLQRRLSYVIRSHCEKENGEPHMDNKIIITSASNYVAIGSKKAY 394
Db 233 VSATFEPDVKSEFLTTHKIDLVLCRAHQVBDEYEFADROLVTVPASNPYCGEFDNAGAM 292
Oy 395 IRLNQLMPHFYGYISAASOTK 416
Db 293 MSVDENIMCSFOYLKPAAKKTK 314

RESULT 29
T25993
phosphoprotein phosphatase (EC 3.1.3.16) 1 ZK354.9 [similarity] - Caenorhabditis elegans:
C.Species: Caenorhabditis elegans
C.Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 05-Oct-2004
C.Accession: T25993
R.Johnson, D.; Wamsley, P.; Bradshaw, H.
submitted to the EMBL Data Library, February 1997
A.Description: The sequence of C. elegans cosmid ZK354.
A.Reference number: Z20120
A.Accession: T25993
A.Status: translated from GB/EMBL/DDBB
A.Molecule type: DNA
A.Residues: 1-310 <JOB>
A.Cross-references: UNIPROT:P91569; UNIPARC:UPI000007FC70; EMBL:U88172; PIDN:AAB42261.1
A.Experimental source: strain Bristol N2; clone ZK354
C.Gene: CBSP.ZK354.9
A.Map position: 4
A.Introns: 27/2
C.Superfamily: serine/threonine protein phosphatase; phosphoesterase core homology; phos-
C.Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-spec
F.15-291/Domin: phosphoprotein phosphatase homology <PPC>
F.42-110/Domin: phosphoesterase core homology <PEC>
F.48_50_76/Binding site: iron (Asp, His, Asp) #status predicted
F.76_108_158_250/Binding site: zinc (Asp, Asn, His, His) #status predicted
F.79_109_274/Active site: Asp, His, Tyr #status predicted
F.80_223/Binding site: substrate phosphate (Arg) #status predicted

Query Match      12.2%; Score 418.5; DB 1; Length 310;
Best Local Similarity 32.2%; Pred. No. 4.9e-19;
Matches 103; Conservative 56; Mismatches 126; Indels 35; Gaps 10;

Oy 98 LPIRNKHIDLLIVPFKKRGNRLHPKYVALILREAAKSLLQPL--NISPVSTAVSQOVTV 155
Db 1 WPILSRPKTIYSQFAEK-----KFI-----KKFNOLNIIFKQPMILIAASPINI 45
Oy 156 CGDLHGKLDLVLVHLKNGLPSSSNPYFVNGDFVDGRKGGLVLLLISLYLAFPNAVPL 215
Db 46 CGDHIHQPSDLLRLPRKNGFRPHAN-YLFLGDYVDRGHGCLSETITLLRAYKIYFNHFFM 104
Oy 216 NRGNHSDSVNNARYGFIREVESKYPRNMKRILLAFIDEYYRWMLPGSVINSRVLIYHG-- 273

```

Db 105 LAGNHESGLINKLIVFYDVECCRRY--NKPSTVYSESGVFSMDPLTALVQGLILCHMGYS 162

Qy 274 -----FSQSTSLDL-ISKIDRGKTVSIIRPPLTGDGEPLDKTEWQOIFDIMWSDP-QATMG 326

Db 163 KQVSNFYDSVQGININIFLOLQWVSGQLAIKRPFPDNEPWTLA-----IDILMSDFTNPQK 217

Qy 327 CVPTNLRGAGWFGGDVTDNPLQRRRLSYVIRSHCKCKNGHGFPHDKILITTFPSASNTYA 386

Db 218 WNPNS-RGVSYVFGSDALRKLIDRLQIDLVRRAHQVDQGYGFAPRNLVTTFAPFYCG 276

Qy 387 IGSNKGAYIRLNNOLMPHFV 406

Db 277 QPDNAAAVMYNNKLVCSFV 296

RESULT 30

S32595

Phosphoprotein phosphatase (BC 3.1.3.16) 1 catalytic chain - yeast (*Saccharomyces cerevisiae*)

N|Alternate names: chromosome disjoining protein Dis2S1; protein YER133w

C|Species: *Saccharomyces cerevisiae*

C|Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text\_change 09-Jul-2004

C|Accession: S32595; S30854; S50636; E32550

R|Feng, Z.; Wilson, S.E.; Peng, Z.Y.; Schlander, K.K.; Reisman, E.M.; Trumbly, R.J. J. Biol. Chem. 266, 23796-23801, 1991

A|Title: The yeast GUC7 gene required for glycogen accumulation encodes a type 1 protein

A|Reference number: S32595; MUID:92084672; PMID:1606885

A|Accession: S32595

A|Status: nucleic acid sequence not shown; translation not shown

A|Molecule type: DNA

A|Residues: 1-312 <FEN>

A|Cross-references: UNIPROT:P32598; UNIPARC:UPI0000052DD3; EMBL:M77175; NID:g172218; PIRTE

A|Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1992

R|Muligan, J.T.; Dietrich, F.S.; Hennessy, K.M.; Sehl, P.; Komp, C.; Wei, Y.; Taylor, submitted to the EMBL Data Library, February 1993

A|Reference number: S30812

A|Accession: S30854

A|Molecule type: DNA

A|Residues: 1-312 <MTU>

A|Cross-references: UNIPARC:UPI0000052DD3; GB:U18916; EMBL:L11120; NID:g1384128; PIDN:AA

A|Title: The sequence of *S. cerevisiae* cosmid 9781, 8198, 9115, 9981, and lambda

A|Description: The sequence of *S. cerevisiae* cosmid 9781, 8198, 9115, 9981, and lambda

A|Reference number: S50636

A|Accession: S50636

A|Molecule type: DNA

A|Residues: 1-312 <DIR>

A|Cross-references: UNIPARC:UPI0000052DD3; EMBL:U18916; NID:g1384128; PIDN:AA

R|Ohura, H.; Kinoshita, N.; Miyatani, S.; Toda, T.; Yanagida, M. Cell 57, 997-1007, 1989

A|Title: The *fls1000* gene required for chromosome disjoining encodes one c

A|Reference number: A32550; MUID:89288305; PMID:2544298

A|Accession: E32550

A|Molecule type: DNA

A|Residues: 1-272, 'V', 274-312 <CHK>

A|Cross-references: UNIPARC:UPI0000175976; GB:M27070

C|Genetics:

A|Gene: SGD:GUC7; DIS2S1

A|Cross-references: SGD:S0000935; MIPS:YER133w

A|Map position: 5R

A|Introns: 59/3

C|Function:

A|Description: phosphoric monoester hydrolase; serine/threonine-specific phosphatase

C|Superfamily: serine/threonine protein phosphatase; phosphoesterase core homology; phosphoesterase core homology; phosphoric monoester hydrolase; serine/threonine-specific phosphatase

C|Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-specific phosphatase

F|57-125/Domain: phosphoprotein phosphatase homology <PFC>

F|63-65, 91/Binding site: iron (Asp, His, Asp) #status predicted

F|91, 123, 172, 247/Binding site: zinc (Asp, Asn, His, His) #status predicted

F|94, 124, 271/Active site: Asp, His, Tyr #status predicted

F|95, 220/Binding site: substrate phosphate (Arg) #status predicted

Query Match 12.2% Score 418.5 DB 2 Length 312

Best Local Similarity 31.5% Pred. No. 4.9e-19

```

Matches 100/ Conservative 63/ Mismatches 125/ Indels 29/ Gaps 9/

QY 103 NHIDLLIDVFRKKRGN--LHPKVALILREAAK-SLKQLPNISPVSTAVSOQVTCGLD 159
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 10 NIIDRLLEVGSKKQGVLDLEBNEIRLGCRAKASIFKQ-----PILBLEAPIKICGI 64

QY 160 HGKIDLLVLAHKNGLPSSNPVFNQDPYDRGKGLFVLLLSLYLAFPAVFLNRGN 219
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 65 HGOYTDLLRLEFYGGFPPEBNN-YLPLGDYVDGKQSLFTICLLAYKIKYKPEFFILRGN 123

QY 220 HEDVWNAARGFIREVSKYPRNHKRIILAFIDEYRNLPLGQSVLNSVLLVHGGS--DST 278
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 124 HECASIRIRYIFTEDECKRY--NIKMKITTD-CFNCPLPAALITDKIPCHGGLSPDLN 180

QY 279 SLDLIKSIDRGKVVSLRPPLTDSBPDKTEWMOQIFDIMSDPOATGVCVPTLRGAGV 338
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 181 SMEDIRRVNR-----PTDIDVGLCDLMSDDPKDIOVGSEBNDGVSFT 225

QY 339 FGPDVTNDFLQRHRLSVIRSHCKPENGHEFMDNKITITFSASNTYAIGSNKGAIVRLN 398
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 226 FGPDVVRFLQKQDMELICRAHGVDEGYEFFGSKQLVTLFSAIPNYGGEFDNAGAMMSVD 285

QY 399 NQLMHFVQYIISAASQT 415
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 286 ESLICSF-QILKPAQKS 301

RESULT 31
S24264
phosphoprotein phosphatase (EC 3.1.3.16) 1A catalytic chain - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C/Accession: S24264
R/Ferreira, P.C.G.; Hemery, A.S.; van Montagu, M.; Inze, D.
submitted to the EMBL Data Library, May 1992
A/Description: Molecular characterization of protein phosphatases 1A from Arabidopsis.
A/Reference number: S24264
A/Accession: S24264
A/Molecule type: mRNA
A/Residues: 1-312 <FER>
A/Cross-references: UNIPROT:P48482; UNIPARC:UP10000160BB8; EMBL:Z12163; NID:g16432; PINDN:
C/Superfamily: serine/threonine protein phosphatase; phosphoserine core homology; phospho-
C/Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-specific
F/36-295/Domain: phosphoprotein phosphatase homology <PP>
F/60-132/Domain: phosphoserine core homology <PSC>
F/70-72.98/Binding site: iron (Asp, His, Asp) #status predicted
F/98-130.179.254/Binding site: zinc (Asp, Asn, His, His) #status predicted
F/101,131,278/Active site: Asp, His, Tyr #status predicted
F/102,227/Binding site: substrate phosphate (Arg) #status predicted

Query Match 12.2% Score 418.5; DB 2; Length 312;
Best Local Similarity 30.3%; Pred. No. 4,9e-19;
Matches 108; Conservative 63; Mismatches 129; Indels 57; Gaps 10;

QY 68 QYQSAHVSVDKDDLVSEFGDIVNAKIELPIRKNHIDLLIDVFRKKG-----NRLL 121
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 4 QGQSMPPALDD-----IIR-----LLDVRNPKPTQKAMLNES 40

QY 122 PKYVALILREAAKSLKQLPNISPVSTAVSOQVTCGLHKKDDLLVLAHKNGLPSSNP 181
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 41 IRQCIYSRI--FLQQPMILIEADIK---ICGIIHQYSDLLRLFEYGGFPPTAN 92

QY 182 YFVNGDVPDVGKQGLFVLLLSLYLAFPAVFLNRGNHDSVWNAARGFIREVSKYPR 241
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 93 YLPLGADVVDGKQSLFTICLLAYKIKYKPEFFILRNCASINRYGGYDECKRF-- 150

QY 242 NHRKIILAFIDEVYKMLPLGQSVLNSRVLLVHGGS--DSTSLDLIKSIDRGKVVSLRP 300
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 151 -SVYLMKVFVDFSGFCVVAAVITDDKILCMHGAISPLDTNVAQIKNIR----- 197

QY 301 DGEPLDKTEWQQLFDIMSDPOATGVCVPTLRGAGVFGPDVTNDFLQRHRLSVIRSH 360
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 198 ---FTDVPDGGGLCDLMSDPSKDVKGMGNNDGVSTFPDPAKAEFLINDMDLICRAH 254

```

QY 361 ECKPENGHEFMHDKITITFSASNTYAIGSKAYITLNNQMLPHFYIISAASQTR 417  
 Db 255 QVVEDGYEFPADKQVLVTITFSAPNYCGEFDNAGAMMSVDSLMCSF-QILKPADKRR 310

## RESULT 32

T09547  
 phosphoprotein phosphatase (EC 3.1.3.16) 1, catalytic gamma chain - alfalfa  
 C/Species: Medicago sativa (alfalfa)  
 C/Date: 16-Jul-1999 #sequence\_rev15on 16-Jul-1999 #text\_change 09-Jul-2004  
 C/Accession: T09547  
 R/Vlast, B.; Goetzda, Toth, B.; Kovacs, I.; Magyar, Z.; Horvath, G.; Bagossi, P.; Gergely, Arch. Biochem. Biophys. 360, 206-214, 1998  
 A/Title: Protein phosphatase 1 catalytic subunit isoforms from alfalfa: Biochemical char  
 A/Reference number: Z16730; MID:99068922; PMID:9851832  
 A/Accession: T09547  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1325 <V1S>  
 A/Cross-references: UNIPROT:O65845; UNIPARC:UP1000009DAB2; EMBL:AJ002486; NID:G3176073;  
 A/Experimental source: subspecies sativa, strain RA3  
 C/Genetics:  
 A/Gene: ppi gamma  
 C/Function:  
 A/Description: catalyzes hydrolysis of peptidyl-phosphoserine or -phosphothreonine to re  
 C/Superfamily: serine/threonine protein phosphatase; phosphoesterase core homology; phos  
 C/Keywords: phosphoric monoester hydrolase; serine/threonine-specific phosphatase  
 F/25-284/Domain: phosphoprotein phosphatase homology <PPP>  
 F/53-121/Domain: phosphoesterase core homology <PEC>

Query Match 12.2%; Score 418.5; DB 2; Length 325;  
 Best Local Similarity 30.1%; Pred. No. 5.2e-19;

Matches 102; Conservative 66; Mismatches 120; Indels 51; Gaps 9;

QY 90 DIVNAKTELPTRKNIHDLIDVPRKKGRNLRHPKYVALILREAKSILKOLPSS----- 143  
 Db 7 DIIRLRIEY-----RNRPG-----KQVVL-----SSEIRQLCNVSRDIFLK 43  
 144 -PVSTAVSQVTVCGDLHGKLDLVLVLRKNGLPSSSNPVFNGDFVDRKGRLEVLTL 202  
 44 QPFLDLQLEAPIKCGVHGVSDDLRLFEYGFPPRAN-YLPIGVDVDRKQSLERITCL 102  
 QY 203 LSLIYLPANVFLNRGNHDSVNNARVGTREVSYPNNKRIILAFIDEVYRLPGSV 262  
 Db 103 LAKIKYPERFLLRGHBCASINRIYGFYDECKRRF---NVLRLMTFTDCSNCLPVAA 159  
 QY 263 LNSRVLIVHGGS-DETSIDLKSIDRGKVSILRPLDGBPLDKTEWQIDIMWSDP 321  
 Db 160 IDKILICMGGLSPDLHNLQIRNLQR-----PTDVPDTGLCDLLMSDP 204  
 QY 322 QAATMGCVPTLRGAGVFGPDVTDFLQRRRLSYVIRSHCKENGHEFMHDKITITFSA 381  
 Db 205 SKVQGMNDKRVSTFGADKITRPLEKIDLDLICRAHGVDEGYEFPANRQVLITISA 264  
 QY 382 SNTYAIGSKAYITLNNQMLPHFYIISAASQTRISP 420  
 Db 265 PNYCGEFDNAGAMMSVDETLMCSF-QILKPADKRRK 302

## RESULT 33

T18936  
 phosphoprotein phosphatase (EC 3.1.3.16) 1-beta catalytic chain - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 15-Oct-1999 #sequence\_rev15on 15-Oct-1999 #text\_change 09-Jul-2004  
 C/Accession: T18936; T25398  
 R/McMurray, A.  
 submitted to the EMBL Data Library, June 1996  
 A/Reference number: Z19047  
 A/Accession: T18936  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-401 <W1S>

A/Cross-references: UNIPROT:O18148; UNIPARC:UP100000787B1; EMBL:Z73968; PIDN:CAA98230.1.  
 A/Experimental source: clone C05A2  
 R/Kelly, P.  
 submitted to the EMBL Data Library, June 1996  
 A/Reference number: Z20028  
 A/Accession: T25398

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA  
 A/Residues: 1-401 <W12>

A/Cross-references: UNIPARC:UP100000787B1; EMBL:Z73977; PIDN:CAA98291.1; GSPDB:GN00023;  
 A/Experimental source: clone T28B11  
 C/Genetics:

A/Gene: CESP:C05A2.1

A/Map position: 5  
 A/Intons: 47/2; 120/1; 159/3; 197/1; 351/3

C/Superfamily: serine/threonine protein phosphatase; phosphoesterase core homology; phos  
 C/Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-spec

F/87-346/Domain: phosphoprotein phosphatase homology <PPP>  
 F/115-183/Domain: phosphoesterase core homology <PEC>

F/121,123,149/Binding site: iron (Asp, His, Asp) #status predicted  
 F/149,181,230,305/Binding site: zinc (Asp, Asn, His, His) #status predicted

F/152,230,312/Active site: Asp, His, Tyr #status predicted  
 F/153,278/Binding site: substrate phosphate (Arg) #status predicted

Query Match 12.0%; Score 413.5; DB 2; Length 401;  
 Best Local Similarity 27.7%; Pred. No. 1.4e-18;

Matches 110; Conservative 74; Mismatches 158; Indels 55; Gaps 12;

QY 61 QAAGRKNQOGSAHVSVDKDLVVEFPDVIYAKTELPTRKNIHDLIDVPRKKGRNRL 120  
 Db 52 QMMRRKNSSSFGELNI-----DQISTVSVVPEMHTLDTVERI----- 92  
 121 HPRYVALILREAKSILKQIPNISPVSTAVSQVTVCGDLHGKLDLVLVLRKNGLPSSSN 180  
 Db 93 -RMVCLARQIFMQPMLIRLEP-----PLKIGDIGHGQPADLRLFNLAGYPPESN 143  
 QY 181 PYFVNDPDRKGRGLEVLTLISLYLAPNNAVFLNRGNHDSVNNARVGTREVSQTP 240  
 Db 144 -YFLGDDVDRPKSLETIVLLCYKIKYNNPFLRGHBEVANLRIYGFYDECKRRY- 201  
 241 RNRKRIILAFIDEVYRLPGSVLNSRVLLVHGGSFDS-TSLDLKSIDRGKVSILRPL 299  
 Db 202 --SVKLMKCFQDVFNCMPAALIDNKIFCCGGLSNLMSLDQKLSR----- 248  
 QY 300 TGBEPLDKTEWQIDIMWSDPQAT-MGCVPTLRGAGVFGPDVTDFLQRRRLSYVIR 358  
 Db 249 ----PCDVQETGLCDVMSDPDATVGVNAPR-REVSVFQVDVLAQFLQKMDLIVR 303  
 QY 359 SHECKPENGHEFMHDKITITFSASNTYAIGSKAYITLNNQMLPHFYIISAASQ---T 415  
 Db 304 GHQVVEDGYEFPGRRLVTVFSAFNYCGEFDNAGAMVNDENILCSF-QILKPSQLVMD 362  
 QY 416 KRLSPKQRMGIYESSALKEIAVRMRDRDELDEPKP 452  
 Db 363 AALAEKQNNFVANTLGMANIKIKQKP-QPMIDKRRFR 398

## RESULT 34

PAEFLA  
 phosphoprotein phosphatase (EC 3.1.3.16) 1-alpha-2 catalytic chain - fruit fly (Drosophi  
 N/Alternate names: phosphoprotein phosphatase isoform 87b catalytic chain  
 C/Species: Drosophila melanogaster  
 C/Date: 31-Mar-1992 #sequence\_rev15on 31-Mar-1992 #text\_change 05-Oct-2004  
 C/Accession: S12960; S05524  
 R/Bombradt, V.; Axton, J.M.; Barker, H.M.; Cohen, P.T.W.  
 FBS Lett. 275, 39-43, 1990  
 A/Title: Protein phosphatase 1 activity in Drosophila mutants with abnormalities in mltc  
 A/Reference number: S12960; MID:91085574; PMID:2175717  
 A/Accession: S12960  
 A/Molecule type: DNA  
 A/Residues: 1-302 <DOM>  
 A/Cross-references: UNIPROT:P12962; UNIPARC:UP1000012488F; GB:X55198; GB:S47852; NID:98;  
 R/Bombradt, V.; Axton, J.M.; Glover, D.M.; Cohen, P.T.W.

Eur. J. Biochem. 183, 603-610, 1989  
 A>Title: Cloning and chromosomal localization of Drosophila cDNA encoding the catalytic  
 A'Reference number: S05524; MUID:8937827; PMID:2550221  
 A'Accession: S05524  
 A'Molecule type: mRNA  
 A'ResIdues: 1-302 <DOM2>  
 A'Cross-references: UNIPARC:UPI000012488F; EMBL:X15583; NID:98363; PIDD:CAA33609.1; PID:  
 C:Genetic:  
 A'Gene: FlyBase:Rpi-87B  
 A'Cross-references: FlyBase:Rpn0004103  
 A'Map position: 3R 87B-12  
 C'Superfamily: serine/threonine protein phosphatase; phosphoserine core homology; phos  
 C'Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-speci  
 F:28-287/Domain: phosphoprotein phosphatase homology <PPP>  
 F:56-124/Domain: phosphoserine core homology <PPC>  
 F:62,64,90/Binding site: iron (Asp, His, Asp) #status predicted  
 F:90,122,171,246/Binding site: zinc (Asp, Asn, His, His) #status predicted  
 F:93,123,270/Active site: Asp, His, Tyr #status predicted  
 F:94,219/Binding site: substrate phosphate (Arg) #status predicted  
  
 Query Match 12.0%; Score 413; DB 1; Length 302;  
 Best Local Similarity 32.3%; Pred. No. 1e-18;  
 Matches 108; Conservative 60; Mismatches 128; Indels 38; Gaps 11;  
  
 QY 89 GIVNAKILPIKKNHIDLLIDVFRKRGKGRNLPKYVALI---LREAKSLKQLPNISPV 145  
 DB 2 GDVNN-----IDSLISRLLEVRGAR-PGRNVQLSGBEIRLCLKSRKIFLSQPI 49  
 QY 146 STAVSQQVTCVGDHKGKLDLLVLHKNGLPSSSNPYVNGDFVDRGKRLVLLLSL 205  
 DB 50 LLELAPLKLICGDIHQYVDLRLFRYGGFPPESSN-YLFLGIVYDRGKSLFTICLLAY 108  
 QY 206 YLAFPAVFLNNGNHDSDVMNARYGIRIVESKYPNNHRIILAFIDRYRMPLSGVLNS 265  
 DB 109 KIRISNPLKRNHNCASINRIYGFYDCKRKY--SILMKITFD-CFNCPLPVAIVDE 165  
 QY 266 RVLIVHGFSS-DTSLDLKSIDRGKYVILRPPLTDSPLDKTEWQIIFDIWMSDP-QA 323  
 DB 166 KIFCGHGLSPDLTSEQRIRIKR-----PTVPPDGLICDLIMSDPDKO 210  
 QY 324 TWGCVNTLRGAGVFGPVDYTNFLQHRRLSYVIRSHCKPNGHEFMHNKIITIFSASN 383  
 DB 211 TWGMEEND-RGVSFTEGAELVAKFLQKHEPDLICRAHQVBDGYEFAPRMLVTLFSASN 269  
 QY 384 YVAGSNKAYIRLNQMLPHFPQYVISAASQTRK 417  
 DB 270 YCGEFNDAGAMMSVDDTLWCSP-QILKPADKRRK 302  
  
 RESULT 35  
 S26225  
 phosphoprotein phosphatase (EC 3.1.3.16) 1 - wild cabbage  
 N'Alternate names: serine/threonine phosphatase type 1  
 C'Species: Brassica oleracea (wild cabbage)  
 C'Date: 03-May-1994 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004  
 C'Accession: S26225  
 R'Bundle: S.J.; Naeffallah, J.B.  
 Plant Mol. Biol. 20, 367-375, 1992  
 A>Title: Molecular characterization of type 1 serine/threonine phosphatases from Brassic  
 A'Reference number: S26225; MUID:93043027; PMID:1330067  
 A'Accession: S26225  
 A'Molecule type: mRNA  
 A'ResIdues: 1-316 <RUN>  
 A'Cross-references: UNIPROT:P48487; UNIPARC:UPI0000131FBB; EMBL:X63558; NID:917881; PIDD:  
 C'Superfamily: serine/threonine protein phosphatase; phosphoserine core homology; phos  
 C'Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-speci  
 F:45-303/Domain: phosphoprotein phosphatase homology <PPP>  
 F:73-141/Domain: phosphoserine core homology <PPC>  
 F:79,81,107/Binding site: iron (Asp, His, Asp) #status predicted  
 F:107,139,187,262/Binding site: zinc (Asp, Asn, His, His) #status predicted  
 F:110,140,286/Active site: Asp, His, Tyr #status predicted  
 F:111,235/Binding site: substrate phosphate (Arg) #status predicted

Query Match 12.0%; Score 412; DB 2; Length 316;  
 Best Local Similarity 30.6%; Pred. No. 1.3e-18;  
 Matches 100; Conservative 61; Mismatches 112; Indels 54; Gaps 10;  
  
 QY 90 DIVNAKILPIKKNHIDLLIDVFRKCR---GNRLHPKVALILREAKSLKQLPNIS--- 143  
 DB 24 DIIRLVLE-----FRNTRGSGKQVH-----LSGG--EIRQLCAVSDI 60  
 QY 144 ----PVSTAVSQQVTCVGDHKGKLDLLVLHKNGLPSSSNPYVNGDFVDRGKRLVLL 199  
 DB 61 FLQQLILLELAPILICGDIHQYSDLRLFRYGGFPPESSN-YLFLGIVYDRGKSLFTI 119  
 QY 200 LLLSLYLAFPAVFLNNGNHDSDVMNARYGIRIVESKYPNNHRIILAFIDRYRMPL 259  
 DB 120 CFLLAKYKIRYENFLLKRNHNCASINRIYGFYDCKRKY--NRLMKITFDICNCLPV 176  
 QY 260 GSVLNSRVLIVHGFSS-DTSLDLKSIDRGKYVILRPPLTDSPLDKTEWQIIFDIW 318  
 DB 177 ALALD-RILCMHGGISPLMSLDQIRSLSR-----PLDIPDSGLVCDLLM 220  
 QY 319 SDPQATWCVNTLRGAGVFGPVDYTNFLQHRRLSYVIRSHCKPNGHEFMHNKIIT 378  
 DB 221 SDPSGDVKGWGANDRGVETFGADTVAAEFLQKNDMLICRAHQVBDGYEFAPRQVTV 280  
 QY 379 PFSASYVIAIGSNKAYIRLNQMLPHF 405  
 DB 281 PFAPNYCEPFDNAGAMMSIDSLWCSP 307  
  
 RESULT 36  
 PABY3  
 phosphoprotein phosphatase (EC 3.1.3.16) PPH3 - yeast (Saccharomyces cerevisiae)  
 N'Alternate names: protein D4421; protein YD8554.08; protein YD8075w  
 C'Species: Saccharomyces cerevisiae  
 C'Date: 30-Jun-1993 #sequence\_revision 10-Feb-1995 #text\_change 05-Oct-2004  
 C'Accession: S44331; S48762; S55822; S49830; C41525; S67892; S17483  
 R'Hoffmann, R.; Jung, S.; Ehrmann, M.; Hofer, H.W.  
 Yeast 10, 567-578, 1994  
 A>Title: The Saccharomyces cerevisiae gene PPH3 encodes a protein phosphatase with prope  
 A'Reference number: S44331; MUID:95028155; PMID:7941742  
 A'Accession: S44331  
 A'Status: not compared with conceptual translation  
 A'Molecule type: DNA  
 A'ResIdues: 1-308 <HOR>  
 A'Cross-references: UNIPROT:P32345; UNIPARC:UPI0000053121  
 R'Coester, F.; Jomiaux, J.L.; Goffeau, A.  
 submitted to the EMBL Data Library, October 1994  
 A'Reference number: S48758  
 A'Accession: S48758  
 A'Molecule type: DNA  
 A'ResIdues: 1-308 <COS>  
 A'Cross-references: UNIPARC:UPI0000053121; EMBL:X82086; NID:9558241; PIDD:CAA57602.1; PI  
 R'Coester, F.; Jomiaux, J.L.; Goffeau, A.  
 Yeast 11, 673-679, 1995  
 A>Title: Analysis of a 32.8 kb segment of yeast chromosome IV reveals 21 open reading fr  
 A'Reference number: S55822; MUID:96093910; PMID:7483840  
 A'Accession: S55822  
 A'Molecule type: DNA  
 A'ResIdues: 1-308 <COM>  
 A'Cross-references: UNIPARC:UPI0000053121; EMBL:X82086; NID:9558241; PIDD:CAA57602.1; PI  
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1994  
 R'Richards, C.; Harris, D.E.  
 submitted to the EMBL Data Library, November 1994  
 A'Reference number: S49823  
 A'Accession: S49830  
 A'Molecule type: DNA  
 A'ResIdues: 1-308 <RIC>  
 A'Cross-references: UNIPARC:UPI0000053121; EMBL:Z46796; NID:9577794; PIDD:CAA6797.1; PI  
 R'Rome, H.; Carlbary, M.; Hu, G.Z.; Nehlin, J.O.  
 Mol. Cell. Biol. 11, 4876-4884, 1991  
 A>Title: Protein phosphatase 2A in Saccharomyces cerevisiae: effects on cell growth and  
 A'Reference number: A41525; MUID:92017761; PMID:1656215

C:Keywords: glycosogen metabolism; heterodimer; iron; metalloprotein; phosphoric monoester  
F:29-288/Domain: phosphoprotein phosphate homology <PP>  
F:75-126/Domain: phosphoesterase core homology <PEC>  
F:63,65,91/Binding site: Iron (Asp, His, Asp) #status predicted  
F:91,123,172,247/Binding site: zinc (Asp, Asn, His, His) #status predicted  
F:94,124,211/Active site: Asp, His, Tyr #status predicted  
F:95,220/Binding site: substrate phosphate (Arg) #status predicted

Query Match            12.0%; Score 411.5; DB 1; Length 333;  
Best Local Similarity 31.0%; Pred. No. 1.5e-18;  
Matches 100; Conservative 65; Mismatches 133; Indels 25; Gaps 8;

Dy         98 LPIRKHHIDLIDVFRKGRNRLHPKYVALI---LRPAASLQQLPISPVSAVSQQVT 154  
        ::::|:::|||||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::  
Dy         1 MDVEKKMLDNMIIRLLREVRGSK-PGVNVOLTESBIKLCKSRERIFLSOELILEAPLK 59  
        |:::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|  
Qy         155 VCGDLHKGTLDDLVLVKRNLPLSSSNPYVNGGFVDNRGGKGLVELLLLILYLAFPNAVF 214  
        ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|  
Db         60 ICDGVHQYYLDLRFYFGFPPEPSN-YLPGLDYDRGKSLEITICILLAYKIYPENPF 118  
        ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|  
Qy         215 LRGNHNEDSVMARRYGFIREVESKYPNNHRLLIAFLDEVVYRWLPGLGSIVNSRVLYVGCF 274  
        ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|  
Db         119 LRGNHECASINRIYGFECKRRY-NIKLMKTFTD-CRNCLPVAIIIDEKIPFCGHGL 175  
        ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|  
Qy         275 S-DSTSLDIKSIDRGKVYSILARPLTDSGPLDKTEWQOLFDMSPPOATMGCVPTWR 333  
        ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|  
Db         176 SPLQSSEQRRTMR-----PTVPDPGLLCBIMSDPDWDVWGGENDR 220  
        |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|  
Qy         334 GAGWFPGEDYTNDPLORHLRSYVIRSHECKPKNGHBPHDKKITTPSASNYYAIGSNKA 393  
        ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|  
Db         221 GVSTTFEPFYAAEFLKHNDLDLICRAHVVEDGEIEFAAKQVLTLSAPNYCGBFDNAGS 280  
        ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|  
Dy         394 YIRLNQMMPHFVOYISAASOTK 416  
        ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|  
Db         281 MMTVDETLMCSF-QILKPADKKK 302  
        |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 58  
T19701  
phosphoprotein phosphatase (BC 3.1.3.16) C34Cl2.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 05-Oct-2004  
C:Accession: T19701  
R,Kershaw, J.  
submitted to the EMBL Data Library, December 1994  
A:Reference number: Z19166  
A:Accession: T19701  
A>Status: translated from GB/EMBL/DDBT  
A:Molecule type: DNA  
A:Residues: 1-382 <'WT'>  
A:Cross-references: UNIPARC:UPI0000172815; EMBL:Z64996; PIDN:CAA87100.1; GSDB:GNO0021;  
A:Experimental source: clone C34Cl2  
C:Genetics:

A/Gene: CBSP.C34Cl2.3  
A/Map position: 3  
A:Intons: 77/2; 107/2; 145/1; 179/2; 221/3; 265/3; 292/3; 336/1  
C/Superfamily: serine/threonine protein phosphatases; phosphoserinase core homology; phos-  
C/Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-spec  
F:96-356/Domain: phosphoprotein phosphate homology <PPP>  
F:124-192/Domain: phosphoesterase core homology <PEC>  
F:130,132,158/Binding site: Iron (Asp, His, Asp) #status predicted  
F:158,190,240,315/Binding site: zinc (Asp, Asn, His, His) #status predicted  
F:161,191,339/Active site: Asp, His, Tyr #status predicted  
F:162,288/Binding site: substrate phosphate (Arg) #status predicted

Query Match            11.9%; Score 409.5; DB 1; Length 382;  
Best Local Similarly 37.3%; Pred. No. 2.4e-18;  
Matches 100; Conservative 45; Mismatches 98; Indels 25; Gaps 7;

Dy         141 NISPVSAAVSGQVTCGDHLGXLDLVLVHKNGLPSSSNPYVNGGFVDNRGGKGLVEL 200  
        |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|  
Db         117 NVNPVSS----PTTGCDIHGGFYDLL-ELFKYGGIVPRNTKVFPMGDYVDRGHSLETVT 171



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: January 20, 2006, 19:38:07 ; Search time 163 Seconds  
(without alignments)

2861.069 Million cell updates/sec

Title: US-09-463-733-1

Perfect score: 3442

Sequence: 1 MDENAIRAIFIOKMYRRHQ.....VEHDIDPTDCSKVIDPKKS 661

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Uniprot\_05.80.\*  
1: uniprot\_sprot.\*  
2: uniprot\_tramb1.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3442	100.0	661	1 RDGC_DROME	P40421 drosophila
2	3121	90.7	603	2 Q5UID3_DROME	Q5UID3 drosophila
3	1930.5	52.1	604	2 Q7Q229_ANOGA	Q7Q229 anopheles g
4	1813	52.7	604	2 Q7Q210_ANOGA	Q7Q210 anopheles g
5	1371.5	39.8	705	2 Q4SOB1_TETNG	Q4SOB1 tetrodon n
6	1359	39.5	757	1 PPE2_MOUSE	PPE2 mouse
7	1341	39.0	707	2 Q61650_CABER	Q61650 caenorhabdi
8	1330	38.6	753	1 PPE2_HUMAN	PPE2 human
9	1317	38.3	707	2 Q01921_CABER	Q01921 caenorhabdi
10	1285.5	37.3	653	2 Q4R4Y0_MACFA	Q4R4Y0 macaca fasc
11	1271.5	36.9	653	1 PPE1_HUMAN	PPE1 human
12	1263	36.7	683	2 Q9M6F4_FUGRU	Q9M6F4 fugu rubrip
13	1238.5	36.0	671	2 Q4SHR7_TETNG	Q4SHR7 tetrodon n
14	1232	35.8	797	2 Q4RU11_TETNG	Q4RU11 tetrodon n
15	1147	33.3	572	2 Q8MYR2_CABER	Q8MYR2 caenorhabdi
16	839	24.4	405	2 Q4R441_MACFA	Q4R441 macaca fasc
17	627.5	18.2	493	2 Q68BP0_XENTR	Q68BP0 xenopus tro
18	618.5	18.0	897	2 Q4QCM7_XENMA	Q4QCM7 xenopus tro
19	615.5	17.9	492	2 Q42205_XENMA	Q42205 xenopus lae
20	615.5	17.9	493	2 Q6G8P6_CABER	Q6G8P6 xenopus lae
21	614.5	17.9	474	2 Q7Q871_ANOGA	Q7Q871 anopheles g
22	613	17.8	874	2 Q4RFR2_TETNG	Q4RFR2 tetrodon n
23	603.5	17.5	953	2 Q4Y018_PLACH	Q4Y018 plasmodium
24	594.5	17.3	923	2 Q81728_TRYCR	Q81728 trypanosoma
25	582.5	16.9	473	2 Q512P5_ENTHI	Q512P5 entamoeba h
26	581	16.9	496	2 Q9N5E8_CABER	Q9N5E8 caenorhabdi
27	578	16.8	829	2 Q4Z5B3_PLABR	Q4Z5B3 plasmodium
28	577.5	16.8	594	2 Q8MQR3_PLABR	Q8MQR3 plasmodium
29	577.5	16.8	658	2 Q8IDR7_PLAF7	Q8IDR7 plasmodium
30	577	16.8	520	2 Q9VH81_DROME	Q9VH81 drosophila
31	575.5	16.7	525	2 Q5CUB8_CRYHO	Q5CUB8 crypsospori

## ALIGNMENTS

RESULT 1	RDGC_DROME	STANDARD;	PRT;	661 AA.
AC	P40421; Q9VMA4;			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DT	13-SEP-2005 (Rel. 48, Last annotation update)			
DE	Serine/threonine protein phosphatase rdgc (BC 3.1.3.16) (Retinal degeneration C protein).			
GN	Name=rdgc; ORFNames=CG6571;			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RP	NCBIOTIDE SEQUENCE (ISORF B).			
RX	MDLLINE=92263398; PubMed=1316807; DOI=10.1016/0092-8674(92)90330-A;			
RA	Steele F.R., Washburn T., Rieger R., O'Tousa J.B.;			
RT	"Drosophila retinal degeneration C (rdgc) encodes a novel			
RL	serine/threonine protein phosphatase.";			
RL	Cell 69:669-676(1992).			
RN	[2]			
RP	NCBIOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].			
RC	STRAIN=Berkley;			
RX	MDLLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;			
RA	Adams M.D., Celniker S.B., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Ananitsides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Wortman J.R., Yandel M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA	Arril J.F., Abmayyan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,			
RA	Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			
RA	Borkov D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			
RA	de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Doup L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Durbin K.J., Egholm G., Ferraz C., Ferrieres S., Fetschmann W.,			
RA	Foster C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K.,			
RA	Harris N.L., Harvey J., Helt G., Gu Z., Guan P., Harris M.,			
RA	Hartis D., Houston K.A., Howland T.J., Hernandez J.R., Houck J.,			
RA	Hosht N., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,			
RA	Jalili M., Kalish F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA	Kamel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,			
RA	Laeko P., Lai Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,			
RA	Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,			
RA	Markov L., Miliushina N.V., Mobarry C., Morris J., Moshrefi A.,			
RA	Mount S.M., Moy M., Murphy B., Murphy K., Muny D.M., Nelson D.L.,			
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,			
RA	Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,			
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,			

32	574.5	16.7	516	2 Q4USE6_THEN	Q4USE6 theileria a
33	573.5	16.7	534	2 Q962N7_PLAF7	Q962N7 plasmodium
34	570.5	16.6	483	2 Q5W6J3_ORYSA	Q5W6J3 oryza sativ
35	570.5	16.6	485	2 Q8H1H4_LYCSE	Q8H1H4 lycopersico
36	570	16.6	484	2 Q9BPW0_HUMAN	Q9BPW0 homo sapien
37	570	16.6	499	1 PPE5_HUMAN	PPE5 homo sapien
38	570	16.6	499	2 Q53FR0_HUMAN	Q53FR0 homo sapien
39	570	16.6	499	2 Q53XV2_HUMAN	Q53XV2 homo sapien
40	569	16.5	479	2 Q64538_RAT	Q64538 rattus norv
41	569	16.5	499	1 PPE5_MOUSE	PPE5 mus musculu
42	569	16.5	499	1 PPE5_RAT	PPE5 rattus norv
43	569	16.5	499	2 Q68G16_RAT	Q68G16 rattus norv
44	568	16.5	499	2 Q5R8T2_PONY	Q5R8T2 pongo pygma
45	567.5	16.5	492	2 Q4PIK6_HUMAN	Q4PIK6 homo sapien



RA Shue B.C., Siden-Kiamos I., Simpson M., Skupek M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Swirskas R., Tector C., Turner R., Venter B., Wang A.H., Wang X.,  
Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*.";  
RT Science 287:2185-2195(2000).  
RN [3]  
RP GENOME REANNOTATION, AND ALTERNATIVE SPLICING.  
RX MEDLINE:24246069; PubMed:12537572;  
RA Miera S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochman S.B.,  
RA Smith C.D., Tully J.L., Whitfield B.J., Bayraktaroglu L., Berman B.P.,  
RA Bellencourt B.R., Gelniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
Lewis S.E.;  
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
RT systematic review.";  
RT Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
CC -1- FUNCTION: Phosphatase required to prevent light-induced retinal  
CC degeneration.  
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +  
CC phosphate.  
CC -1- COFACTOR: Binds 1 iron ion per subunit (By similarity).  
CC -1- COFACTOR: Binds 1 manganese ion per subunit (By similarity).  
CC -1- ENZYME REGULATION: Regulated by calcium.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Name=B; Synonyms=C;  
CC IsoId=P40421-1; Sequence=Displayed;  
CC Name=A;  
CC IsoId=P40421-2; Sequence=VSP\_009324;  
CC Note=No experimental confirmation available;  
CC -1- TISSUE SPECIFICITY: Expressed in the visual systems of the fly, as  
CC well as in the mushroom bodies of the central brain.  
CC -1- SIMILARITY: Belongs to the PPP phosphatase family.  
CC -1- SIMILARITY: Contains 3 EF-hand domains.  
CC -1- SIMILARITY: Contains 1 IQ domain.  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
DR EMBL: M86628; AAB00734.1; -; Genomic DNA.  
DR EMBL: AE003514; AAF49044.2; -; Genomic DNA.  
DR EMBL: AE003514; AAO41217.1; -; Genomic DNA.  
DR PIR: A42287; A42287.  
DR HSSP: Q08209; IAU1.  
DR Ensemble: CG5571; *Drosophila melanogaster*.  
DR FlyBase: FBgn004366; rdcG.  
DR GO: GO:000516; Focalmodulin binding; IDA.  
DR GO: GO:0019722; P-calcium-mediated signaling; TAS.  
DR GO: GO:0016059; P-deactivation of rhodopsin mediated signaling; IMP.  
DR GO: GO:0006470; P-protein amino acid dephosphorylation; IMP.  
DR InterPro: IPR011992; EF-hand type.  
DR InterPro: IPR002048; EF-hand\_Ca\_Bd.  
DR InterPro: IPR000048; IQ\_CaM\_Bd\_region.  
DR InterPro: IPR004843; M-pesterase.  
DR InterPro: IPR012008; PPRF.  
DR InterPro: IPR006186; T\_phlase\_aph.  
DR Pfam: PF000936; ehand; 3.  
DR Pfam: PF00612; IQ; 1.  
DR Pfam: PF00149; Metallophos; 1.  
DR PIRSF: PIRSF00912; PPRF; 1.  
DR PRINTS: PR00114; STRPHATS.  
DR ProDom: PD000012; EF-hand; 1.

DR ProDom: PD000252; T\_phlase\_aph; 1.  
DR PROSITE: PS00018; EF\_HAND\_1; 2.  
DR PROSITE: PS02222; EF\_HAND\_2; 3.  
DR PROSITE: PS50096; IQ; 1.  
DR PROSITE: PS00125; SER\_THR\_PHOSPHATASE; 1.  
KW Alternative splicing; Calcium; Hydrolyase; Iron; Manganese; Manganese;  
KW Metal-binding; Protein phosphatase; Repeat; Sensory transduction;  
KW Vision.  
FT DOMAIN 7 32 IQ.  
FT DOMAIN 441 476 EF-hand 1.  
FT DOMAIN 526 561 EF-hand 2.  
FT DOMAIN 566 601 EF-hand 3.  
FT CA\_BIND 539 550 1 (Potential).  
FT CA\_BIND 579 590 2 (Potential).  
FT REGION 105 413 Catalytic.  
FT ACT\_SITE 220 220 Proton donor (By similarity).  
FT METAL 158 158 Iron (By similarity).  
FT METAL 160 160 Iron (By similarity).  
FT METAL 187 187 Iron and manganese (By similarity).  
FT METAL 219 219 Manganese (By similarity).  
FT METAL 271 271 Manganese (By similarity).  
FT METAL 360 360 Manganese (By similarity).  
FT VARSPPLIC 1 58 Missing (in isoform A).  
FT FTId=VSP\_009324.  
SQ SEQUENCE 661 AA; 75511 MW; A3DC42933B4CCA33 CRC64;  
Query Match 100.0%; Score 3442; DB 1; Length 661;  
Best Local Similarity 100.0%; Pred. No. 2e-200;  
Matches 661; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MDENAIRAIFQKRYRHOARREMOORCNNOIPONTLYASBODAEYKPPNDLIKMP 60  
1 |=====|  
Db 1 MDENAIRAIFQKRYRHOARREMOORCNNOIPONTLYASBODAEYKPPNDLIKMP 60  
QY 61 QAAGRKNOYQSAHYSVLDDKDDVEEFQDYVNAXIEPIRKNHIDLIDVPRKRGRL 120  
61 |=====|  
Db 61 QAAGRKNOYQSAHYSVLDDKDDVEEFQDYVNAXIEPIRKNHIDLIDVPRKRGRL 120  
QY 121 HPKTYALILREANGLKQLPNISPVSTAVSQVTCGLHGLDLYLVLRNGLPSSGN 180  
121 |=====|  
Db 121 HPKTYALILREANGLKQLPNISPVSTAVSQVTCGLHGLDLYLVLRNGLPSSGN 180  
QY 121 HPKTYALILREANGLKQLPNISPVSTAVSQVTCGLHGLDLYLVLRNGLPSSGN 180  
121 |=====|  
Db 121 HPKTYALILREANGLKQLPNISPVSTAVSQVTCGLHGLDLYLVLRNGLPSSGN 180  
QY 181 PVEFGDVPDGRGLAVTLTLTLSTLYLAFPNAPVFLNRNHDVSNARYGFRVRSYRP 240  
181 |=====|  
Db 181 PVEFGDVPDGRGLAVTLTLTLSTLYLAFPNAPVFLNRNHDVSNARYGFRVRSYRP 240  
QY 241 RHHKRIILAFIDSVYRWMLPLGSLVNSRVLIYVGGFSDSLDLSIDRGKYVILRPPLT 300  
241 |=====|  
Db 241 RHHKRIILAFIDSVYRWMLPLGSLVNSRVLIYVGGFSDSLDLSIDRGKYVILRPPLT 300  
QY 301 DGEPLDKTEWQIPIIMSDPOATMGCVPTLRGAGVFGPDVTNFIQRRLSYVIRSH 360  
301 |=====|  
Db 301 DGEPLDKTEWQIPIIMSDPOATMGCVPTLRGAGVFGPDVTNFIQRRLSYVIRSH 360  
QY 361 ECKPVGHEPMNDKTIITPSASNYVAIGSNQAVYRLNNOLMPHVQYISAASQYRISF 420  
361 |=====|  
Db 361 ECKPVGHEPMNDKTIITPSASNYVAIGSNQAVYRLNNOLMPHVQYISAASQYRISF 420  
QY 421 KQRMQIVSSALKEIAYMRDRDELDEFRYPDKDGGYISISMKCYMENVTKLGLPW 480  
421 |=====|  
Db 421 KQRMQIVSSALKEIAYMRDRDELDEFRYPDKDGGYISISMKCYMENVTKLGLPW 480  
QY 481 RILRDKLAPGDSQKVNTRTLDLDTVILAEADQMSVMDALYANTASIVAINITIDA 540  
481 |=====|  
Db 481 RILRDKLAPGDSQKVNTRTLDLDTVILAEADQMSVMDALYANTASIVAINITIDA 540  
QY 541 DNSGRIITDEFTALDILVAHNPQAYSKAEMLCKRMMDLNDGDVNLNFIARLBDL 600  
541 |=====|  
Db 541 DNSGRIITDEFTALDILVAHNPQAYSKAEMLCKRMMDLNDGDVNLNFIARLBDL 600  
QY 601 HRKEQDENIRRSRSGPSVAKTADPTYLADKISKNTLVYEHIDPTDCSKYIDPPK 660  
601 |=====|  
Db 601 HRKEQDENIRRSRSGPSVAKTADPTYLADKISKNTLVYEHIDPTDCSKYIDPPK 660



```

Qy      661 S 661
Db      661 S 661

RESULT 2
OSUID3_DROME PRELIMINARY; PRT; 603 AA.
AC
OSUID3;
DT 01-FEB-2005 (TEMBLrel. 29, Created)
DT 01-FEB-2005 (TEMBLrel. 29, Last sequence update)
DE 01-FEB-2005 (TEMBLrel. 29, Last annotation update)
DE RH46370p.
GN Name=rdgc;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxId=7227;
OX
RN
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Berkeley;
RA Stapleton M., Carlson J., Chavez C., Friese B., George R., Pachtel J.,
RA Park S., Wan K., Yu C., Rubin G.M., Celniker S.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; BT015959; AAV36844.1; -, mRNA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR004843; M-pesterase.
DR InterPro; IPR006186; T_phthase_apah.
DR Pfam; PF00036; ehnd; 3.
DR Pfam; PF00149; Metallophos; 1.
DR PRINTS; PR00114; STPHPTASR.
DR ProDom; PD000012; EF-hand; 1.
DR ProDom; PD000252; T_phthase_apah; 1.
DR SMART; SM00054; Efh; 3.
DR SMART; SM00156; PF2AC; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN_2.
DR PROSITE; PS00125; SER_THR_PHOSPHATASE; UNKNOWN_1.
SQ SEQUENCE 603 AA; 68121 MW; A2FBCA6A32CE2403 CRC64;

Query Match 90.7%; Score 3121; DB 2; Length 603;
Best Local Similarity 99.8%; Pred. No. 5,3e-181;
Matches 602; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      59 MPOAARKKQYQGSAAHVSVLDDKDDLYEERGDIVNAKIELPRKKNHIDLLIDVFRKKRN 118
Db      1 MPOAARKKQYQGSAAHVSVLDDKDDLYEERGDIVNAKIELPRKKNHIDLLIDVFRKKRN 60

Qy      119 RLHPKTVALLREAAKSLKQLPNISPVSTAVSQQVTVCGDLHGKLDLVLVLRKGLPSS 178
Db      61 RLHPKTVALLRRAAKSLKQLPNISPVSTAVSQQVTVCGDLHGKLDLVLVLRKGLPSS 120

Qy      179 SNPYVNGDFVDRGKGLLEVLLLSLYLAFPNAVFLNRGNHDSVMAARYGEFIREVSK 238
Db      121 SNPYVNGDFVDRGKGLLEVLLLSLYLAFPNAVFLNRGNHDSVMAARYGEFIREVSK 180

Qy      239 YPNHNRILLAFIDEVYRWLPLGSVLSNRVLIYVGGSDSTSLDISIRGKVSILRPP 298
Db      181 YPNHNRILLAFIDEVYRWLPLGSVLSNRVLIYVGGSDSTSLDISIRGKVSILRPP 240

Qy      299 LIDGEPIDKTEWQOIFDIMMSDPAATMGCVPTLRAGAVFGPDVTNPLQRHRLSYVR 358
Db      241 LIDGEPIDKTEWQOIFDIMMSDPAATMGCVPTLRAGAVFGPDVTNPLQRHRLSYVR 300

Qy      359 SHECKNGHEFMADNKITITFSASNTYAGSNKGAVIRLNNOMLPHVOYISAAQTKL 418
Db      301 SHECKNGHEFMADNKITITFSASNTYAGSNKGAVIRLNNOMLPHVOYISAAQTKL 360

Qy      419 SPQQRNGIVSSALKELAVMRDHRDELDEPRKYVPOKSGYISISHWCKMKNENTKGL 478

```

```

Db      361 SPQQRNGIVSSALKELAVMRDHRDELDEPRKYVPOKSGYISISHWCKMKNENTKGL 420
Qy      479 PWRLLRDKLAPGTDQKVNYNRTLLDITDVLLEAEDGMSVMDALYANKASIVAFNII 538
Db      421 PWRLLRDKLAPGTDQKVNYNRTLLDITDVLLEAEDGMSVMDALYANKASIVAFNII 480

Qy      539 DADNNGEITLDEPERTIDLLVAMHPGAYSKAEMLKCRMMNDLNGDKVLEPFLRIS 598
Db      481 DADNNGEITLDEPERTIDLLVAMHPGAYSKAEMLKCRMMNDLNGDKVLEPFLRIS 540

Qy      599 DLHREKQDENIRRSSTGRPSVAKTADPTVTLADKISKNTLVESHIDPTDCESKYIP 658
Db      541 DLHREKQDENIRRSSTGRPSVAKTADPTVTLADKISKNTLVESHIDPTDCESKYIP 600

Qy      659 KXS 661
Db      601 KXS 603

RESULT 3
ID 07Q229 ANOGA
ID 07Q229 ANOGA PRELIMINARY; PRT; 613 AA.
AC 07Q229;
DT 01-MAR-2004 (TEMBLrel. 26, Created)
DT 01-MAR-2004 (TEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DR ENSANGP000004936 (Fragment).
GN ORFNames=ENSANG00000003837;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophellinae; Anopheles.
OX NCBI_TaxId=180454;
RN
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RA The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RA The Anopheles gambiae re-annotation.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RA The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008966; RA12933.2; -, Genomic DNA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR004843; M-pesterase.
DR InterPro; IPR006186; T_phthase_apah.
DR Pfam; PF00036; ehnd; 2.
DR Pfam; PF00612; IQ; 1.
DR Pfam; PF00149; Metallophos; 1.
DR PRINTS; PR00114; STPHPTASR.
DR ProDom; PD000012; EF-hand; 1.
DR ProDom; PD000252; T_phthase_apah; 1.
DR SMART; SM00054; Efh; 2.
DR SMART; SM00156; PF2AC; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN_2.
DR PROSITE; PS00096; IQ; 1.
DR PROSITE; PS00125; SER_THR_PHOSPHATASE; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 613
SQ SEQUENCE 613 AA; 69803 MW; BA0AABAEPF8D548 CRC64;

Query Match 56.1%; Score 1930.5; DB 2; Length 613;
Best Local Similarity 59.6%; Pred. No. 1e-108;
Matches 366; Conservative 101; Mismatches 130; Indels 17; Gaps 6;

```

Qy	5	ALTAATFICQWYRRHARREBORCMQIFQULFVASBODAEUYKFPNDIKMPAAG	64
Db	4	AMRAAILIQWYRRFLARIEIRRRYTWTFQSEIETAGSDQVRLYFNFNALHTIPETAG	63
Qy	65	RKQYQGSAAHVSLVDKDDLVEBFDIVNAKIELPIRKNHIDLIDVERKKRGRLHPKY	1244
Db	64	RPLDSQNTSRSSSAADEGILTPERG-YQGPDIKFPIDKKELEVIDLFRKKK-NRLHAKY	1211
Qy	125	VALLIEPAASLQLQPNISPVSTAVSGQYTVCGDILHGKLDLVLVHNGLPSSSNPYF	1844
Db	122	VAGILRAATSKLRLRPNLNQASPAISKQYTIIGDHLGKLDLVLVFNKNGLPSPENPYF	1811
Qy	185	NGDFVDRGKGLSVLLLSLVLAFPNVAFLRGNHDSVMNARYGFRVYSKYPNNHK	2444
Db	182	NGDFVDRGKGLSVLLLSLVLAFTLVFPQGVFLRGNHEDVMNARYGFRVYHQYKINAB	2411
Qy	245	RLIAFLDEVYRWLPGLGSVLNSRYLVHGFSDSTSLDIKSIDRGKYSILRPLTD---	3010
Db	242	RLKLIDEVYRWLPGLGTIVNNRVLVHGISIDSTDLIRSLDRKYSILRPLTPETSTA	3010
Qy	302	--GEPIDKTEWQOIFDIMSDPOLATMGCVPTLIRGAGVFGPDVDTNLFQHRHLSYIRS	3599
Db	302	PGAEIIDKYEWMQVPIILMSDDPHTEGCRPNELRAGTYFGPDVTSKFLQRYKQYLVRS	3619
Qy	360	HECKPNHEHMHNKIITTPASANNYATISNGAYIRLANQLMHFPQYISAAQTRLG	4199
Db	362	HECKPDHEHMHGKXITTPASANNYEIGSNNGAVALDPLDTHFHWYTAASAKYKLT	4219
Qy	420	FKPMGIVESSALKELAVMRDHRDLDEBFKYPCKOSGYISISHMKVMENTYKGLP	4799
Db	422	FRQVGVYESSALRELAAKLRBRRIELEBFSPRPKNKYVLPLAKNCALRGSTISGLP	4819
Qy	480	WRLIRDKLAPGTS-----QKVNYNRTLDLDDTVILREADGMSVMDALYANKAST	5319
Db	482	WRMLKCLAPADNPSTAVTAVTVESYRKTIHLDDTDFKSNQNTTSVAESLYKNSSL	5419
Qy	532	VAFNIITDADNSEITLDEFETAILDLVHMHGAYSKAMLEKRMNDLNDGKVDLNEF	5919
Db	542	EAFLRILDKNSQISLEBFGACELRRHFP-NHTHSQLLDMCRMDINDKGVLNBER	6009
Qy	592	LEAFRLSDLHRKEQ 605	
Db	601	LETFRLCE-NAKEQ 613	

RESULT 4  
Q7ON10\_ANOGA  
ID O7ON10\_ANOGA PRELIMINARY; PRT; 604 AA.

AC Q7ON10;  
DT 01-MAR-2004 (TrEMBLrel). 26, Created)  
DT 01-MAR-2004 (TrEMBLrel). 26, last sequence update)  
DT 01-MAR-2004 (TrEMBLrel). 26, last annotation update)  
DE ENSANGP0000000059 (Fragment).  
GN ORFNames=ENSANG0000000057;  
OS Anopheles gambiae str. PEST.  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;  
OC Anophelinae; Anopheles.  
OX NCBI\_TaxId=180454;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=PEST;  
RG The Anopheles gambiae Sequence Committee;  
RT "Anopheles gambiae re-annotation".  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=PEST;  
RG The Anopheles gambiae Sequence Committee;  
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
CC -!- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.

[illegible]

ID Q4SOB1 TETNG PRELIMINARY; PRT; 705 AA.  
AC Q4SOB1;  
DT 13-SEP-2005 (TREMUR-rel. 31, Created)  
DT 13-SEP-2005 (TREMUR-rel. 31, Last sequence update)  
DT 13-SEP-2005 (TREMUR-rel. 31, Last annotation update)  
DE Chromosome 4 SCAF14533, whole genome shotgun sequence.  
DN (Fragment).  
GN ORFNames=GSTENG00014444001;  
OS Tetraodon nigrovittatus (Green puffer).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthopterygii; Acanthopterygii; Perciformes; Tetraodontiformes;  
OC Tetraodontidae; Tetraodontidae; Tetraodon.  
OX NCBI\_TaxID=99883;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,  
RA Mauceli E., Bounneau L., Fischer C., Ozouf-Costraz C., Bernot A.,  
RA Nicoud S., Jaffe D., Fisher S., Lutfalla G., Dosat C., Segurens B.,  
RA Dastila C., Salanoubat M., Levy M., Boudet N., Castellano S.,  
RA Anthonard V., Jubin C., Castell V., Katinka M., Vacherie B.,  
RA Cnaud C., Duprat S., Brotilier P., Coutanceau J.P., Gouzy J.,  
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McSwan P., Bosak S.,  
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mestrov J.,  
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,  
RA Lander E., Schacher V., Quetier F., Saurin W., Scarpelli C.,  
RA Wincker P., Lander E.S., Weissbach J., Roest Crollius H.,  
RA "Genome duplication in the teleost fish Tetraodon nigrovittatus reveals  
RT the early vertebrate proto-karyotype."  
RL Nature 431:946-957(2004).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RG Genoscope, Whitehead Institute Centre for Genome Research;  
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
CC -1- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL; CAAB01014533; CAP97171.1; -; Genomic\_DNA.  
FT NON\_TER 1  
FT NON\_TER 1  
SQ SEQUENCE 705 AA; 81541 MW; B4513BDBA03B1FOA CRC64;  
Query Match 39.8%; Score 1371.5; DB 2; Length 705;  
Best Local Similarity 39.2%; Pred. No. 1e-74;  
Matches 278; Conservative 117; Mismatches 186; Indels 129; Gaps 10;  
QY 5 AIRPAIFIOKRYRHOARRRORRCMQRIFONLEVASBODQALYKPFNDLIGMFOAG 64  
DB 9 AIRPALLIQRMVQYARLEMRRTYTNIPQSTIEVGEODQIKYDPPFASLMERFPAS 68  
QY 65 RKN-----OYQGANVSVLDDKDLVEEFGDVIWAKIELPIRKNHID 106  
DB 69 ERNLISHIFPKNEVCCEADWERFYSVSEVPEPSS-----GPRTPFLITCGVL 118  
QY 107 LIDVRRKRGKRNLRHYVALIIRPAKSLKQLPNISPVSTAVSQOYTCGLHGLDDI 166  
DB 119 KIVEAKFKHQ-QLHARYVLQLGERTWRPFLRILPININQVSACONRRIITCGDLHGLDDI 176  
QY 167 LVVLHK-----NGLPSSNPVYVNGDFDRGKRGLEVLILLSLYLAPNAPVFNRRN 219  
DB 177 LVVYRGELVPAONGPSSSEKPEVFGDFDRKSSIEILLFGLVLPVNDVHLNRGN 236  
QY 220 HEDSVNARVGYRREVSRYPRNKKILAFIDEVYMLPLGVLNRSVLVHGSPDSTS 279  
DB 237 HEDHIVLRVGFYKRWGKRVHGKILKLGKIFSLPLATLVHDKVILVHGISDITD 296  
QY 280 LDIKSIDRGKYSILRPP----- 299  
DB 297 LDIARVDRHKYVSALRPPLKQANAKRAGAAGPWEGRRRVYTLTYKSTADHSHKRP 356  
QY 300 -----TDGSP-----LDKTEMQ 311

DB 357 RSHLSQQLSGGVNMEBELKKRRRLAGFDQSYGEPQSRSDSDSPESGEATEDEHMK 416  
QY 312 QTPDIWSPDQATMGCVPTTLGAGVWFGPDYTDNFIQHRRLSYRSHCKPNGEFPAH 371  
DB 417 QVVDLWSPMPQSGCIPNEVGGGCGYCPDYEKTLASHDDIQLVRSHECKQEGYFCH 476  
QY 372 DKIIITIPASNYVAIGSNKGYIRLNQIMPHVQYISAAGQTKLSKORNGYIYSSA 431  
DB 477 DRRLVITFSASNTVEGSRNGAVIRKGDILIPHFVY-QASIKCRKLTLLKQSGRTESA 535  
QY 432 LKELAVRMDHDEDEBRKXDPKDSGYISISHWKVMEVYKGLPRLRLDKLAPGT 491  
DB 536 LRLAKQLPVHSDLSMAFQEPDPNNTGMSLRHMASATERVLNGLPFRVLRPQVLSST 595  
QY 492 DSGKNTYRFTL-DLDTVILPABADGMSVMDALYANKASLVAFNIIDADNSGITLDE 550  
DB 596 ESGMDVYQQWIRPEFSTTERKLEH-ADN-SILEMYRNHNLLETIFPIIDTDSGLISPEB 653  
QY 551 PRTAIDLVAAHPGAYSKAMLEKCRMDNDGKVDLNEPLRAPLSDL 600  
DB 654 FHOYKLLSSHLKMEISDKAVTDLAQSIDFNKDSIDINEFMEAFRLVHL 703  
RESULT 6  
PBR2\_MOUSE  
ID\_PBR2\_MOUSE STANDARD; PRT; 757 AA.  
AC Q35385;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Serine/threonine protein phosphatase with EF-hands-2 (EC 3.1.3.16)  
DN (PBR-2).  
GN Name=Pbr2;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridae; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Eye;  
RX MEDLINE=97471020; PubMed=9326663; DOI=10.1073/pnas.94.21.11639;  
RA Sherman P.M., Sun H., Macke J.P., Williams J., Smallwood P.M.,  
RA Nathans J.;  
RT "Identification and characterization of a conserved family of protein  
RT serine/threonine phosphatases homologous to Drosophila retinal  
RT degeneration C";  
RL Proc. Natl. Acad. Sci. U.S.A. 94:11639-11644(1997).  
RN [2]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
RC STRAIN=C57BL/6; TISSUE=Retina;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins P.S., Wagner L., Shennan C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.P., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loughran N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McKernan K.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butcherfield J.S.N., Krzywicki M.T., Skaleka U., Smalins D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -1- FUNCTION: May play a role in photoreduction. May

dephosphorylate photoactivated rhodopsin. May function as a calcium sensing regulator of ionic currents, energy production or synaptic transmission.

-1 CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein + phosphate.

-1 COFACTOR: Binds 1 iron ion per subunit (By similarity).

-1 COFACTOR: Binds 1 manganese ion per subunit (By similarity).

-1 ENZYME REGULATION: Activated by calcium (By similarity).

-1 TISSUE SPECIFICITY: Detected in retina, more specifically in photoreceptors.

-1 SIMILARITY: Belongs to the PPP phosphatase family.

-1 SIMILARITY: Contains 3 EF-hand domains.

-1 SIMILARITY: Contains 1 IQ domain.

-----

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

CC

DR EMBL, AF023458, AAB82798.1, -, mRNA.

DR EMBL, BC027049, AAH27049.1, -, mRNA.

DR HSSP, P09860, 1DTL.

DR Ensembl, ENSMUSG00000029410, Mus musculus.

DR MGI, MGI:1342304, Ppaf2.

DR InterPro, IPR011592, EF-Hand type.

DR InterPro, IPR02048, EF-hand\_Ca\_bd.

DR InterPro, IPR000048, IQ\_Cam\_Bd\_region.

DR InterPro, IPR004843, M-Pesterase.

DR InterPro, IPR012008, PPSF.

DR InterPro, IPR006186, T\_phase\_apah.

DR Pfam, PF00036, etehand, 3.

DR Pfam, PF00612, IQ, 1.

DR Pfam, PF00149, Metallophos, 1.

DR PIRSF, PIRSF000912, PPSF, 1.

DR PRINTS, PRS0114, STPHPHASE.

DR Prodom, PP000012, EF-hand, 1.

DR Prodom, PP000252, T\_phase\_apah, 1.

DR SMART, SM00054, Efh, 3.

DR SMART, SM00156, Pp2ac, 1.

DR SMART, SM00015, IQ, 1.

DR PROSITE, PS00018, EF\_HAND\_1, 3.

DR PROSITE, PS00222, EF\_HAND\_2, 3.

DR PROSITE, PS50096, IQ, 1.

DR PROSITE, PS00125, SSR\_THR\_PHOSPHATASE, 1.

KW Calcium; Hydrolase; Iron; Manganese; Metal-Binding; Protein phosphatase; Repeat; Sensory transduction; Vision.

KV

FT DOMAIN 21 46

FT DOMAIN 572 607

FT DOMAIN 656 691

FT DOMAIN 696 731

FT CA\_BIND 669 680

FT CA\_BIND 709 720

FT REGION 128 544

FT ACT\_SITE 241 241

FT METAL 179 179

FT METAL 181 181

FT METAL 208 208

FT METAL 240 240

FT METAL 292 292

FT METAL 492 492

SEQUENCE 757 AA; 86645 MW; 821B3D1061AC0C8 CRC64;

Query Match 39.5%; Score 1359; DB 1; Length 757;

Best Local Similarity 38.9%; Pred. No. 6.5e-74;

Matches 278; Conservative 113; Mismatches 200; Indels 124; Gaps 7;

3 ENAIRPAALFIQKWRHQAARERQRCNMQJFONLEYSBODQALYKFNLLIKENQA 62

17 EKAFPAQYALAIQWRYRVARALERMRRCNMFNFQSIETVAGQDDQVLTAFHFPIVDHPTPS 76

63 AGKKQVQGSAAHVSLVDKDLVIEFGDI-----VNAKIELPIKKNHIDLLIDVER 113

Db	77	SHHEEDFLNRK#TERFADVETBEGDFES1EVPDSTYGPLSLPFLDLPHATALVEAR	136
Qy	114	KKGRRLHPKVALILIREAAKSLKQLPNISPVSTAVSQOVTYCGDILHGKLDLVLVLRN	173
Db	137	LRQ--QLHARVYANLLETRFKHLAQDLPNINRSTCVSBEVTVCGDILHQGLDLDLIF	194
Qy	174	GLPSSSNFYVNGDVYDNGKRGLEVLILLISLYLAFPNNAVFLNRGNHDSVMNARYG	233
Db	195	GLPSPERAYVFNPGDVPDVRGKDSVEVLTMTVFAFMLYVPKSFHLNRGNHSHL	254
Qy	234	EYESYCPANHKKRIILAFIDVYRWMLPLGSLNERNVYLIVHGGSFDSYSLDIKSID	293
Db	255	EVMKRYKIHGKKILLTDQVFCMLPLATLVDEKVLVHOGVSDKTDLSILAKIDRHK	314
Qy	294	ILR-----	296
Db	315	TMRCTKRSENEREQKKNDNTSSGQKPTWFLQPSRSLPSSPFLISGFAFYKAGRC	374
Qy	297	-----PLPTD-----GEP	304
Db	375	SIPCGSPNSKELSRGQVRRSVLDLEQCRQAGFLGIREKESLPLADADCVADGGCV	434
Qy	305	LDKT--EMQOQLFDIMWSDPOATMVGVPMTLRGAGVFGPDVYDNLFGRRLSVTR	362
Db	435	LEPTEERKQVVDILMSDPAQBGCKAANAIRGQGCFFGVDVIERLMEKQLQULLRSH	494
Qy	363	KPNGEHFEMDNKIITIFASNYAIAIGSNKGAIRLNNQMLPFPVOYISAASTKRLS	422
Db	495	KPEGEGFCHNRKVLITIFASNYIEVGSNRGAIVKLGPAITPIIVQY-QANKATRL	553
Qy	423	RMGVIYSSALXELAVRMDHNDLEDEFRKYDPKDSGYISISHCKVMENVTGLPWR	482
Db	554	RISRVEBALRALRQKLPFASHSDDLIVERKRPDPDSGVIITLSDMTAVESVHLGL	613
Qy	483	LRDLKAPETDQKYVNRRTLDLDDTVLLBAAQDSVWDALYANKASVALFNITDAON	542
Db	614	LRPOLVNSSANVLEYRSMVLSLAKEQ-LSRNNIGSSILEKLYRNRSNLETFRIIDSD	672
Qy	543	SGEITLDFEFYALDILVAHMPGAYKAEMLTECRMMNDLNGDKVDLNEFLAERL	597
Db	673	SGFISLDLFRQTKLFSHMSIDITDDGICDLAKSIDRKKQGHIDINEFLAERL	727

DR Pfam; PF00149; Metallophos; 1.  
 DR PIRSF; PIRSF000912; PPERF. 1.  
 DR PRINTS; PRO00114; STEPHHTSR.  
 DR ProDom; PD000012; EF-hand; 1.  
 DR ProDom; PD000252; T\_phthase\_aph; 1.  
 DR SMART; SM00054; EPh; 2.  
 DR SMART; SM00156; PR2Ac; 1.  
 DR PROSITE; PS00018; EF\_HAND; UNKNOWN\_2.  
 DR PROSITE; PS00125; SER\_THR\_PHOSPHATASE; UNKNOWN\_1.  
 DR Hypothetical protein.  
 DR SEQUENCE 707 AA; 79840 MW; 4ADE2CA1D186C630 CRC64;  
 Query Match 39.0%; Score 1341; DB 2; Length 707;  
 Best local Similarity 42.4%; Pred. No. 7.3e-73;  
 Matches 269; Conservative 118; Mismatches 192; Indels 56; Gaps 12;  
 QY 6 IRAAIFIQKWRHQAARREMRQCNQIFQNTETASBODQALYKFFNDLIKHPQAAG 65  
 DB 83 IKSALIQKWRRCARLEARRATQIFTALEYAEQDQKLYDFADVIRAMABENGK 142  
 QY 66 KNOYQG--SAHVSVDL--KDLVSEFGDIYNAKIF-----LPINK 102  
 DB 143 TGGDNERNSPMLSHYAKPSLMDSEGTVMKMLDTSFGNVVDNRNYKPYLSPLDK 202  
 QY 103 NHIDLIDYFRKRGKRLHPKYVALILREAAKSLKQLPNISPVTASQOVTWQGLHGK 162  
 DB 203 POYAKMTEAF--KMKVLIHPKYVLMILHEARKIFKSPSPVSRISTISNQITTCGLHGK 260  
 QY 163 LDDLVLVHKNGLPSSSNIPVFGDVPDGRKGLLEVLLLSLYLAFPNAVFLRGNHED 222  
 DB 261 PDDLCLILKNGVPSVDNPYIFNGDVPDGGQSEVLCVLFALVIVDPMSIYNRGNHD 320  
 QY 223 SVNNARYGTRIEVSKYTPRNHRIILAFIDETVYMLPLGVSANRVLIVHGFSFSTSLD 282  
 DB 321 HMMNLEYGPIKEISTYKDLSTPTITLLBVFSEMLIATIVDKDIFVVGISDQTEVER 380  
 QY 283 IKSIDRGVYSILRPPLTDG-----BPLDKTEHQQIFDIWNSPOATMGCVPTLHG 334  
 DB 381 LDKIPHRQSVLRKPVNKKMDSAEKENVQSYDEKQMLDIMSPOKNGCWPVVFVG 440  
 QY 335 AGVWGPVDTYDNELQHRILSYVIRSHCKPNCHGFHNDKITTFASNYAIGSKGAY 394  
 DB 441 GGSYFPADITASTLEHGFRLLVIRSHCKEGEGEFHNTCTLVFASNYTEGSKGAY 500  
 QY 395 IRLNQLMHPVQVYISAQSTKRLS-FKQRMGIVESSALKEALVRMRDRHLEDEFRKY 453  
 DB 501 VKFYGSKRHFVQYM--ASKTRKSTLRRLRGVSESAVLEKSKLSFTPLDQKEFENV 558  
 QY 454 DPRDSGYTISISHCKVMENVTKGLPWRLLRDLAEGTSQK-VATNRITLIDLITVYIL 512  
 DB 559 DLKRSGNLPILKMSCEVERITGMLPWIALAPVATLSDGKRVNKEKRIKAQVGTHA 618  
 QY 513 AEAAGSWDAIYANKASLVAIFNIIIDANGSEITLDEFEITADILVAHMPGAYSK---- 568  
 DB 619 QEKD---IYSLYRHKSTLETFLFRMDKNSGVSKKEPFDACEVL-----GKTYRPLQ 670  
 QY 569 ----AEMLEKRMMDLNGDKVDLNFLEAPRLSD 599  
 DB 671 TDYIAQIAES---IDENKQGFIDLNLLEAFRLVD 702  
 RESULT 8  
 ID PPER2\_HUMAN STANDARD; PRT; 753 AA.  
 AC 014830; 014831;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 13-BFP-2005 (Rel. 48, Last annotation update)  
 DE Serine/threonine protein phosphatase with EF-hands-2 (EC 3.1.3.16)  
 DR (PPERF-2).  
 OS Name=PPER2;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;  
 OC Homo  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE (ISOFORMS PPERF-2(S) AND PPERF-2(L)).  
 RC TISSUE=Retina;  
 RX MEDLINE=97471020; PubMed=9326663; DOI=10.1073/pnas.94.21.11639;  
 RA Sherman P.M., Sun H., Macke J.P., Williams J., Smallwood P.M.,  
 RA Nathans J.;  
 RT "Identification and characterization of a conserved family of protein  
 RT serine/threonine phosphatases homologous to Drosophila retinal  
 RT degeneration C.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 94:11639-11644(1997).  
 CC -1- FUNCTION: May play a role in phototransduction. May  
 CC dephosphorylate photoactivated rhodopsin. May function as a  
 CC calcium sensing regulator of ionic currents, energy production or  
 CC synaptic transmission.  
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +  
 CC phosphate.  
 CC -1- COFACTOR: Binds 1 iron ion per subunit (By similarity).  
 CC -1- COFACTOR: Binds 1 manganese ion per subunit (By similarity).  
 CC -1- ENZYME REGULATION: Activated by calcium (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Localized to photoreceptors, PPERF-2(L) is at  
 CC least 2 fold more abundant in rod inner segments than in the outer  
 CC segments.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=PPERF-2(L);  
 CC IsoId=O14830-1; Sequence=Displayed;  
 CC Name=PPERF-2(S);  
 CC IsoId=O14830-2; Sequence=VSP\_005103, VSP\_005104;  
 CC -1- TISSUE SPECIFICITY: Retinal specific (Probable).  
 CC -1- SIMILARITY: Belongs to the PPP phosphatase family.  
 CC -1- SIMILARITY: Contains 3 EF-hand domains.  
 CC -1- SIMILARITY: Contains 1 IQ domain.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 CC EMBL; AF023456; AAB82796.1; -; mRNA.  
 CC EMBL; AF023457; AAB82797.1; -; mRNA.  
 CC DR HSSP; P09860; IDTL.  
 CC DR Ensembl; ENSG00000156194; Homo sapiens.  
 CC DR HGNC; HGNC:9244; PPERF2.  
 CC MIM: 602256; -;  
 CC DR GO; GO:0004722; P:protein serine/threonine phosphatase activity; TAS.  
 CC GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.  
 CC DR InterPro; IPR011992; EF-Hand type.  
 CC DR InterPro; IPR002048; EF hand Ca bd.  
 CC DR InterPro; IPR000048; IQ Cam bd region.  
 CC DR InterPro; IPR004843; M-Pesterase.  
 CC DR InterPro; IPR012008; PPERF.  
 CC DR InterPro; IPR006186; T\_phthase\_aph.  
 CC DR Pfam; PF00036; ehand; 3.  
 CC Pfam; PF00612; IQ; 1.  
 CC Pfam; PF00149; Metallophos; 1.  
 CC PIRSF; PIRSF000912; PPERF. 1.  
 CC PRINTS; PRO00114; STEPHHTSR.  
 CC ProDom; PD000012; EF-hand; 1.  
 CC ProDom; PD000252; T\_phthase\_aph; 1.  
 CC ProDom; PD00054; EPh; 3.  
 CC SMART; SM00015; IQ; 1.  
 CC SMART; SM00156; PR2Ac; 1.  
 CC PROSITE; PS00018; EF\_HAND\_1; 2.  
 CC PROSITE; PS00222; EF\_HAND\_2; 3.  
 CC PROSITE; PS00096; IQ; FALSB\_NEG.  
 CC PROSITE; PS00125; SER\_THR\_PHOSPHATASE; 1.  
 CC KW Alternative splicing; Calcium; Hydrolyase; Iron; Manganese;  
 KW Metal-binding; Polymorphism; Protein phosphatase; Repeat;  
 KW Sensory transduction; Vision.

```

FT DOMAIN 21 46 IQ.
FT DOMAIN 568 603 EF-hand 1.
FT DOMAIN 652 687 EF-hand 2.
FT DOMAIN 692 727 EF-hand 3.
FT CA_BIND 665 676 1 (Potential).
FT CA_BIND 705 716 2 (Potential).
FT REGION 128 540 Catalytic.
FT ACT_SITE 241 241 Proton donor (By similarity).
FT METAL 179 179 Iron (By similarity).
FT METAL 181 181 Iron (By similarity).
FT METAL 208 208 Iron and manganese (By similarity).
FT METAL 240 240 Manganese (By similarity).
FT METAL 292 292 Manganese (By similarity).
FT METAL 488 488 Manganese (By similarity).
FT VARBPLIC 588 598 IRTSDMAAVE -> ASSQLCTYQOK (in isoform
PREF-2(S)).
FT VARBPLIC 599 753 /FTID-VSP_005103.
FT VARBPLIC 120 120 Missing (in isoform PREF-2(S)).
FT VARIANT 120 120 /FTID-VSP_005104.
FT S -> R.
SQ SEQUENCE 753 AA, 86431 MM, 289D4AC240193A0D CRC64,
/FTID-VAR_010230.
Query Match 38.6%; Score 1330; DB 1; Length 753;
Best Local Similarity 38.1%; Pred. No. 3,7e-72;
Matches 278; Conservative 109; Mismatches 185; Indels 158; Gaps 11;
3 ENAIRAIRITQKMYRRHQAARREMRQRCNQIFQNTLEYASBODAEIYKFEFNDLIKMPQA 62
17 ERAFAKAAALIQRYRRYVARLEEMRRRCTSIFQSIYAGQQOQVKLHDFSYLMQH----- 72
63 AGRKNOYQSAHVSVLDDKDDLYEERGD-----IYNAKIEIP 99
73 -----FIRSSH---NDRDFLRITTEBPADDSMKKCSYDESIEVPDSTYGPESLSP 122
100 IRKNHIDLLIDYFRKKRGRNLHPKYVALIRKAAKSLKQLPNISPVSTAVSOQVTCGDL 159
123 LIPDHATALVEAFRLNQ--QLHARYLNLTYERKKLVOLPINNRSTCYSEIITVCGDL 180
160 HGKLDLVLVHKNGLPSSSNPVYFGDVPYDRGKGLVLLLSLYLAFPNVAFNLNGN 219
181 HGQLDLIFIFYKNGGLPSPERSVYFNGDVPYDRGKGSVEILMLIFAPMLVYPKXEFHLNRGN 240
220 HEDSVNARYGFRREVESKYPRNHRLLAFIDSVYRWLGLSVTLNRYLVIVGGPDSYS 279
241 HEDHAWNLAYGFLKXWANKYVHGKELTLDQVFCMLPLATLIDKVLILHGVSVDITD 300
280 LDIKSIDRGKYVSIILR-----PPL- 299
301 LELDLIERSKIYSTMRCKTROKSEKOMEKRRANKQSSAQQPIFWFLPESRSLPSSPLR 360
300 -----TDEGPLDK----- 307
361 LGSYKAQKTRSSSSIPCSGLDRELSRQVRSVLELRCHQOAGLVTGEKEPSSRA 420
308 -----TEMQOIPDIWSDPQATWGCVPNTLRAGVAFPGDVTNPLQRR 352
421 SEADSEAGELRKPTQSEMRQVVDILMSDPAQSGCAANTIRGGGCVFGDVTQOLLQKN 480
353 LSVVIRSHCKPGHGFEMDNKIIITIFSASNYAISNKGAYIRLNNQMLPHVQYISA 412
481 LQFLIRSHCKPGGYFCNRRKVLITIFSASNYEVSNGAYVKGALPALTPHVOY-QAN 539
413 SOTKRLSPKRMGIYESSALKEBLAVWRDRDELDEFRKYPKSGYISISHCKVMEN 472
540 KYTHLTLMQRISLRVSESAIRALREKLFANSSDLSEFFKHADKGLITLSDMAAAVBS 599
473 VTKGLGPMYLLRDKLAPGTDQKVNNTKLLDITVYILEABDGSVNDALYANKASLY 532
600 VHLHSGPMYMLRKQLVNSSADNMLRYKSWLKNLAKQ--LSRENTIGSSLLETTYRNSNLE 658
533 ALFNIIIDANGSGITLDEFEETALIDLVANHPGAYSAEMLEKC-----RWMDLNGDGKVD 587

```

```

DB 659 TIFRIIDSDHSGFSLDESRQTKLFSSSH-----NIDITDDICIDLARSIDFNKGCHID 713
QY 588 INEFLEAPRL 597
DB 714 INEFLEAPRL 723
RESULT 9
001921_CABEL
ID 001921_CABEL PRELIMINARY; PRT; 707 AA.
AC 001921; 017047;
DT 01-JUL-1997 (TRMBLrel. 04, Created)
DT 01-OCT-2001 (TRMBLrel. 18, Last sequence update)
DT 13-SEP-2005 (TRMBLrel. 31, Last annotation update)
DE Phosphatase with ef hands protein 1, isoform a (Protein phosphatase
DE with EF-hands).
GN Name=pef-1; ORFNames=PF23H11.8;
OS Caenorhabditis elegans.
OC Rhabditidae; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditoidea;
OC Rhabditidae; Peloderae; Caenorhabditis.
OX NCBI_TaxID=6239;
(1)
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Br1etol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT Investigating biology."
RL Science 282:2012-2018 (1998).
(2)
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=97471020; PubMed=9326663; DOI=10.1073/pnas.94.21.11639;
RA Sherman P.M., Sun H., Macke J.P., Williams J., Smallwood P.M.,
RA Nathans J.;
RT "Identification and characterization of a conserved family of protein
RT serine/threonine phosphatases homologous to Drosophila retinal
RT degeneration C."
RL Proc. Natl. Acad. Sci. U.S.A. 94:11639-11644 (1997).
DR EMBL, AF003389; AACT7138.2; -; Genomic_DNA.
DR EMBL, AF023454; AAB82794.1; -; mRNA.
DR PIR, T34072; T34072.
DR PIR, T42239; T42239.
DR HSSP, P36873; 1176.
DR Ensembl, F23H11.8; Caenorhabditis elegans.
DR Wormbase, WBGene0003969; pef-1.
DR Wormpep; F23H11.8a; CE27999.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR011992; EF-Hand type.
DR InterPro; IPR002048; EF-hand_Ca_bd.
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR004843; M-Pesterase.
DR InterPro; IPR012008; PPEF.
DR InterPro; IPR006186; T_phase_apah.
DR Pfam; PF00036; ehand; 2.
DR Pfam; PF00612; IQ_1.
DR Pfam; PF00149; Metallophos; 1.
DR PIRSF; PIRSF000912; PPEF_1.
DR PRINTS; PR00114; STEPHATASE.
DR Prodom; PD000012; EF-hand; 1.
DR Prodom; PD000252; T_phase_apah; 1.
DR SMART; SM00054; Efh; 3.
DR SMART; SM00015; IQ_1.
DR SMART; SM00156; PZAC; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN_2.
DR PROSITE; PS00125; SER_THR_PHOSPHATASE; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 707 AA, 80330 MM, 1943CE21BD0A7EF7 CRC64;
Query Match 38.3%; Score 1317; DB 2; Length 707;
Best Local Similarity 42.1%; Pred. No. 2,1e-71;
Matches 266; Conservative 116; Mismatches 200; Indels 50; Gaps 12;

```

QY 6 IRAAIFIQKYYRHHOARREMOQRCCNNOIFONLEYASBODAEIYKFPNDLIKMPQOAGR 65  
 DB 83 IKAIIIOKYYRCEARLEARRATWOFIPIALYASBODOLKIYDFADVIRMAEENK 142  
 QY 66 KNOYQG--SAHVSULD--KDDIVEFGDIYNAKIE-----LPIRK 102  
 DB 143 GGVENGRNSPLMSALSHYAKPSLMSDSGFTVKKMLBETDPTSDVNDIDRNYKPTLSLPIDK 202  
 QY 103 NHTDLLIDVPRKGRKRLHKKYVALLIREAKSLKOLPNISPVSTAVSQOYTCGDLHGK 162  
 DB 203 POUAKKIEAF--RYNRYLHKKYVALLIREAKSLKOLPNISPVSTAVSQOYTCGDLHGK 260  
 QY 163 LDDLVLVHKNGLPSSSNPFVNGDFVDRGKGLLEVLLLSLYLAFPNVAFPLNRGNHD 222  
 DB 261 FDDLCLILYKNGPSPDNPIIFNGDFVDRGKGLLEVLLLSLYLAFPNVAFPLNRGNHD 320  
 QY 223 SVNNARVGFIREVSKYPRNHRKILAFIDEVYRMLPGSVLNSRVLIYHGGFSDSTSLD 282  
 DB 321 HINMLRYGFIKELSTYKXKDLSTPTITRLBEDVFSMLPATIIDRIDFVHGGISDQTEVSK 380  
 QY 283 IKSIDGKYVSLRPPUTDGEPLDK-----TEKQOIFDIMSPOATMGCVPTLRGA 335  
 DB 381 LDKIPHRFQSVLRPPVNNKMESEKNSAVNVDEMKOMLIMWSDKONKGCWPNYFRGG 440  
 QY 336 GWFPGDVTDNFLOHRLSYVIRSHCKPNGHEFMDNKKITIFASANTYATGNNGAYI 395  
 DB 441 GSFPGDITASFLEKNGFRLVNSHCKEYGFESHNTCLTFVSANTYETSSNGAYI 500  
 QY 396 R-LNNOLMPHFVOYIISAASQTKLS--FKQMGIVESSALKEIYVRRDRHDELEDFRKY 453  
 DB 501 KFIQSKQHPFYQV--ASKTHKSLTRELGYVESAVYELKSLSPHTDQKFEIM 558  
 QY 454 DPDSGYSISHWCKVMEVNTKGLPWRLLRDKLAP-GTDSQKVNTNRTLDLDTVILE 512  
 DB 559 DIEKSKLPLKMSDCVERITGMLPWLALAPVATLSEDKYVMEKDRIRIYQVGTIA 618  
 QY 513 ASADGQWMDALYANKASLVAIFNIIDANSGEITLDEPETAIDLVAHMPGYSKAEV 571  
 DB 619 QEKD---IYBSLFRHSTLTETLFRPMKONSGOVSKMEFLDACEVL-----GKYTRPLQ 670  
 QY 572 ----LEKCRMDLNGDGKVDLNEFLBAFRLSD 599  
 DB 671 TDYISQIASIDENKGFIDNBLBAFRLVD 702  
 RESULT 10  
 Q4R4YO MACFA PRELIMINARY: PRT: 653 AA.  
 AC Q4R4YO-  
 DT 13-SEP-2005 (T-EMBLrel. 31, Created)  
 DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)  
 DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)  
 DE Brain cDNA, clone: Q4R4-11374, similar to human protein phosphatase, EF hand calcium-binding domain 1 (PBPFI), transcript variant 1.  
 OS Macaca fascicularis (Cebus eating macaque) (Cynomolgus monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Cercopithecoidea; Cercopithecoidea; Macaca.  
 NCBI\_TaxID=9541;  
 [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA International consortium for macaque cDNA sequencing, analysis, RT "DNA sequences of macaque genes expressed in brain or testis and its evolutionary implications".  
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Oseada N., Hirata M., Tanuma R., Kusuda J., Hida M., Suzuki Y., Sugano S., Gojobori T., Shen J.C.-K., Wu C.I., Hashimoto K., "Substitution rate and structural divergence of 5'UTR evolution: comparative analysis between human and cynomolgus monkey cDNAs." J. Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL, AB169764; BAE01845.1; -, mRNA.

SQ SEQUENCE 653 AA; 75406 MW; A54FCA236554516F CRC64;  
 Query Match 37.3%; Score 1285.5; DB 2; Length 653;  
 Best Local Similarity 40.8%; Pred. No. 1.5e-69;  
 Matches 261; Conservative 122; Mismatches 209; Indels 47; Gaps 9;  
 QY 3 ENAIRAIFIQKYYRHHOARREMOQRCCNNOIFONLEYASBODAEIYKFPNDLIK-----57  
 DB 14 ETSKALILIQWYRKYKARLROHYALTIPQISIBYABOQWOLSTFFSFMEYAHV 73  
 QY 58 HMPQAGRNQYQASAHVAVLDDKDLVEFGDIYNAKIEFLPRKHIDLLIDVPRKGR 117  
 DB 74 HKEBELRQNSLESBODLDRHDYVDLIDVPSYNPRLOPLCTDIDLLLEAFEGQ 132  
 QY 118 NRIHPRTVALIREAKSLKOLPNISPVSTAVSQOYTCGDLHGKLDLVLVHKNGLPS 177  
 DB 133 -ILHAHYLEVLFETKQVLCQMPNPHVQTSPEKELTICGDLHGKLDLFLMRYKGLPS 191  
 QY 178 SGNPVPNGDFVDRGKGLLEVLLLSLYLAFPNVAFPLNRGNHDSVNNARVGFIREVES 237  
 DB 192 ESNPVPNGDFVDRGKNSIRILILICVSPFLVYPNDLHNRGNHEDPMMLRYGFTREILH 251  
 QY 238 KYPNRHRIILAFIDEVYRMLPGSVLNSRVLIYHGGFSDSTSLDIFKSIDGKYVSLR 297  
 DB 252 KYKMGKRLIQLLEBILYAMLPGITVYDNEILVHGGISDITDNLHRYBRNKKSVLLP 311  
 QY 298 PL-----TDGEP---LDKTEKQOIFDIMSPOATMGCV 328  
 DB 312 PIETNRDNDTDSKNNKVGTFNAHGRIKTNGSPSEHLPHHEMQIIDILMSDRGNGGC 371  
 QY 329 PRTLRGAGVPRGPDVTNLOHRLSYVIRSHCKPNGHEFMDNKKITIFASANTYATG 388  
 DB 372 PVTGRGGGCFPGPDVSKILNKYOLKMLIRSHCKEYGFICHGDKVNTIFASANTYEBG 421  
 QY 389 SNKGAITIRLNNOLMPHFVOYIISAASQTKLSF--KQMGIVESSALKEIYVRRDRHDE 445  
 DB 432 SNRGATIKCSGTAPRFPY---QYKATCQPLRQRYTMESSAIKIKERIVIRKSD 487  
 QY 446 LDEPRKPDPKDSGYISISHWCKVMEVNTKGLPWRLLRDKLAPGTDSQKVNTNRTLDL 505  
 DB 488 LTRAFQLOHRRSGKLSQWAFCEMNTLGLNLPMSLSNLVNIDKGNVETMSFQNI 547  
 QY 506 DTDVILIEADGQWMDALYANKASLVAIFNIIDANSGEITLDEPETAIDLVAHMPGA 565  
 DB 548 HIEKPYE--BAHS--TLVETLYRYSDELIFNAIDTDSGLSIEBFRAMMKLFSAHYHYL 605  
 QY 566 YSKAEMLBKCRMDLNGDGKVDLNEFLBAFRLSDLRKE 604  
 DB 606 IDDSOVNKLANTIMDLNKGSDIDPNEFLKAFV--VHYRY 642  
 RESULT 11  
 PPEL HUMAN STANDARD: PRT: 653 AA.  
 AC 014829; 015253; Q9NUZ1; Q9UHO;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 13-SEP-2005 (Rel. 48, Last annotation update)  
 DE Serine/threonine protein phosphatase with EF-hands-1 (BC 3.1.3.16) (PBPFI-1) (protein phosphatase with EF calcium-binding domain) (PBPFI)  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea; Homo.  
 NCBI\_TaxID=9606;  
 RP NUCLEOTIDE SEQUENCE.  
 RA TISSUE=Retina;  
 RX MEDLINE=97471020; PubMed=9326663; DOI=10.1073/pnas.94.21.11639;  
 RA Sherman P.M., Sun H., Macke J.P., Williams J., Smallwood P.M., Nathans J.;



RT "Identification and characterization of a conserved family of protein  
RT serine/threonine phosphatases homologous to Drosophila retinal  
RT degeneration C.";  
RT Proc. Natl. Acad. Sci. U.S.A. 94:11639-11644(1997).  
RL [2]  
RN NUCLEOTIDE SEQUENCE, AND ALTERNATIVE SPLICING.  
RC TISSUE=Fetal brain;  
RC MEDLINE=97358589; PubMed=9215685; DOI=10.1093/hmg/6.7.1137;  
RA Montini E., Rugat E.I., van de Vosse E., Andolfi G., Mariani M.,  
RA Puca A.A., Conales G.G., den Dunnen J.T., Balibio A., Franco B.,  
RT "A novel human serine-threonine phosphatase related to the Drosophila  
RT retinal degeneration C (rdgc) gene is selectively expressed in sensory  
RT neurons of neural crest origin."  
RL Hum. Mol. Genet. 6:1137-1145(1997).  
RN [3]  
RN NUCLEOTIDE SEQUENCE.  
RC TISSUE=Retina;  
RC MEDLINE=98104127; PubMed=9430683; DOI=10.1074/jbc.273.3.1462;  
RA Huang X., Honkaniemi R.E.;  
RT "Molecular cloning, expression, and characterization of a novel human  
RT serine/threonine protein phosphatase, pp7, that is homologous to  
RT Drosophila retinal degeneration C gene product (rdgc)."  
RL J. Biol. Chem. 273:1462-1468(1998).  
RL [4]  
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RN PubMed=15772651; DOI=10.1038/nature03440.  
RA Rose M.T., Graham D.V., Coffey A.J., Scherer S., McElay K., Muzny D.,  
RA Platzer M., Howell G.R., Burrows C., Bird C.P., Frankish A.,  
RA Lovell J.L., Howe K.L., Ashurst J.L., Fulton R.S., Sudbrak R., Wen G.,  
RA Jones M.C., Hume M.E., Andrews T.D., Scott C.E., Seale S.,  
RA Ramser J., Whitlatch A., Deadman R., Carter N.P., Hunt S.E., Chen R.,  
RA Cree A., Gunaratne P., Havlak P., Hodgson A., Metzger M.L.,  
RA Richards S., Scott G., Steffen D., Sodergren E., Wheeler D.A.,  
RA Worley K.C., Alnsouh R., Ambrose K.D., Ansari-Lari M.A., Aradhya S.,  
RA Ashwell R.I., Babbage A.R., Baggsley C.L., Ballabio A., Banerjee R.,  
RA Barker G.E., Barlow K.F., Barrett I.P., Bates K.N., Beare D.M.,  
RA Beasley H., Beasley O., Beck A., Bethel G., Blechschmidt K., Brady N.,  
RA Bray-Allen S., Bridgman A.M., Brown A.J., Brown M.J., Bonin D.,  
RA Bruford E.A., Buhay C., Butch P., Burford D., Burgess J., Buttrill W.,  
RA Burton J., Bye J.M., Carder C., Carrel L., Chako J., Chapman J.C.,  
RA Chavez D., Chen E., Chen G., Chen Y., Chen Z., Chinault C.,  
RA Ciccodicola A., Clark S.Y., Clarke G., Clee C.M., Clegg S.,  
RA Clerc-Blankenburg K., Clifford K., Cobley V., Cole C.G., Conquer J.S.,  
RA Corby N., Connor R.B., David R., Davies J., Davis C., Davis J.,  
RA Delgado O., Dehaene D., Dharm P., Ding Y., Dinh H., Dodsworth S.,  
RA Drepper H., Dugan-Rocha S., Dunn M., Durbin K.J., Dutta I.,  
RA Bades T., Blwood M., Emery-Cohen A., Eyrington H., Evans D.M.,  
RA Faulkner L., Francis P., Frankland J., Fraser A.E., Galoczky P.,  
RA Gilbert J., Gill R., Gloeckner G., Gregory S.G., Gribble S.,  
RA Griffiths C., Grocock R., Gu Y., Gwilliam R., Hamilton C., Hart E.A.,  
RA Hawes A., Heath P.D., Heilmann K., Hennig S., Hernandez J.,  
RA Hinemann B., Ho S., Hofits M., Howden P.J., Huckle R.J., Hume J.,  
RA Hunt P.J., Hunt A.R., Isherwood J., Jacob L., Johnson D., Jones S.,  
RA de Jong P.J., Joseph S.S., Keenan S., Kelly S., Kershaw J.K., Khan Z.,  
RA Kioschke P., Klages R., Knights A.J., Koshiura A., Kovar-Smith C.,  
RA Laird G.K., Langford C., Lawlor S., Leverish M., Lewis L., Liu W.,  
RA Lloyd C., Lloyd D.M., Lonsdale H., Loveland J.E., Lovell J.D.,  
RA Lozadow R., Lu J., Lyne R., Ma J., Maheshwari M., Matthews L.H.,  
RA McDowell J., McLaren S., McMurtry A., Meldal P., Mellinger T.,  
RA Milne S., Miner G., Mistry S.L., Morgan M., Morris S., Mueller J.,  
RA Mullikin J.C., Nguyen N., Nordesiek G., Nyakatura G., O'dell C.N.,  
RA Okunou G., Palmer S., Pandian R., Parker D., Parrish J.,  
RA Pasternak S., Patel D., Pearce A.V., Pearson D.M., Pelan S.E.,  
RA Perez L., Porter K.M., Ramsey Y., Reichwald K., Rhodes S.,  
RA Ridler K.A., Schlessinger D., Schubler M.G., Sehra H.K.,  
RA Shaw-Smith C., Shen H., Sheridan R.M., Showkhan R., Stuce C.D.,  
RA Smith M.L., Socher B.C., Steingrub H.E., Steward C.A., Storey R.,  
RA Swann R.M., Searbeck D., Tabor P.E., Taudien S., Taylor T.,  
RA Teague B., Thomas K., Thorpe R., Timms K., Tracey A., Trevanion S.,  
RA Tromms A.C., d'Urso M., Verdusco D., Villaseña D., Waldron L.,  
RA Wall M., Wang C., Warren G., Warr G.L., Wei X., West A.,  
RA Whitehead S.L., Whitely M.N., Wilkinson J.E., Willey D.L.,  
RA Williams G., Williams L., Williamson A., Williamson H., Wilming L.,  
RA Woodmansey R.L., Wray P.W., Yen J., Zhang J., Zhou J., Zoghbi H.,  
RA Zorilla S., Buck D., Reinhardt R., Pousetta A., Rosenthal A.,  
RA Lehnach H., Weindl A., Minx P.J., Hillier L.W., Willard H.F.,  
RA Wilson R.K., Waterston R.H., Rice C.M., Vaudin M., Coulson A.,  
RA Nelson D.L., Weinstein G., Sulston J.E., Durbin R., Hubbard T.,  
RA Gibbs R.A., Beck S., Rogers J., Bentley D.R.;  
RT "The DNA sequence of the human X chromosome."  
RL Nature 434:325-337(2005).  
RN [5]  
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).  
RN TISSUE=Testis;  
RN PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins P.S., Wagner L., Sherman C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,  
RA Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Hsieh P.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ueda T.B., Toshynski S., Carlini P., Prange C.,  
RA Rahn S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullihy S.J.,  
RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmitt J., Myers R.W.,  
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [6]  
RN SPLICING ISOFORM(S) THAT ARE POTENTIAL NMD TARGET(S).  
RN PubMed=14759258; DOI=10.1186/gb-2004-5-2-r8;  
RA Hillman R.T., Green R.B., Brenner S.R.;  
RT "An unappreciated role for RNA surveillance."  
RL Genome Biol. 5:RESEARCH008.1-RESEARCH008.16(2004).  
CC -1- FUNCTION: May have a role in the recovery or adaptation response  
CC of photoreceptors. May have a role in development.  
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +  
CC phosphate.  
CC -1- COFACTOR: Binds 1 iron ion per subunit (By similarity).  
CC -1- COFACTOR: Binds 1 manganese ion per subunit (By similarity).  
CC -1- ENZYME REGULATION: Activated by calcium.  
CC -1- BIOPHYSICOCHEMICAL PROPERTIES:  
CC pH dependence:  
CC Optimum pH is 8.0;  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=5;  
CC Name=1;  
CC IsoId=O14829-1; Sequence=Displayed;  
CC Name=1A;  
CC IsoId=O14829-2; Sequence=VSP\_005098;  
CC Name=1B;  
CC IsoId=O14829-3; Sequence=VSP\_005099;  
CC Name=2;  
CC IsoId=O14829-4; Sequence=VSP\_005100; VSP\_005101;  
CC Note=May be produced at very low levels due to a premature stop  
CC codon in the mRNA, leading to nonsense-mediated mRNA decay;  
CC Name=3;  
CC IsoId=O14829-5; Sequence=VSP\_005102;  
CC Note=May have no functional significance;  
CC -1- TISSUE SPECIFICITY: Detected in retina and retinal derived Y-79  
CC retinoblastoma cells. Also found in fetal brain.  
CC -1- SIMILARITY: Belongs to the PPP phosphatase family.  
CC -1- SIMILARITY: Contains 3 EF-hand domains.  
CC -----  
CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its





134 Q- --TTHARVVLQHLHETRLKLLKQMPNIIMHSTYYTEKITTCGGDLHGQOLDLLILTYKNG 191

175 LPSSSNPYFVNGDPVDRGRKGLVEVLLLLLSLYLAPFNAYFLNRGNHEDSVNARVGPTR 234

192 LPASLTPYFVNGDPVDRGNKSVVWMLLRAVYLLIPDYVHMLNKGNEHDMNLRVGPTR 251

235 VESKTPRNHKLILAFIDVEYRWMLPLGVLSNRYLVIYHGGSFDSLSIDLKSIDRGYVSI 294

252 VMQYKIHOCSEILQFQDVFSLLPIATIIDKILYIHGGSIDQTDLPDLSIERHNIKSA 311

295 LRPPLTIDG----- 303

312 LRFPTVYQVLSIGSSRQVKIRISTDSSRPSSSRSHNIORTPHGGSVFPQLPFLDLSLA 371

394 ---PLDKTEMQOIFDLMWGDPOATMGCVPTNLRGAVWFGPVDYDNLPHRLRSLVYIRSH 360

372 PSPPHQGMQKQIVDLIMSDPDKLKCCCPNTRFGGCGTGPDPVTRLLHLGHGLIRSH 431

351 ECKPNGHEFMHNDKIITTPSASNYVIAIGSNKQAYIRLNQMLMHPVOYISAAQSOTRLSP 420

432 ECKQGYELCHGQGVITTPISASNYVBEGSNRQAYIVGHEIMRFPQY-QVNSSTKRLT 490

421 KQRMGVESALKELARVRKDRHDELEDFRKTDPDQSGYISISHMKYMENTKGLGLP 480

491 TORVAAESASARLAKELKPTRHSEILGFOEDHNNTGVILVSEWAOVLETRLDLP 550

481 RLRLRKLATGDSOKVNYNRITDLIDVLIKAEADQMGVMDLVANKSLVAFINIIA 540

551 RLRLRHLVRLAGGRVDYQSCFEDMERGV-PQVEVT-PIALDLFRYRDIETITPISIDK 608

541 DNGEITLDEFETALIDLVAHMPGAYSKAMLEKCMMDLNGDKVDLNEFLFAFLSIDL 600

609 DQSGIISTEFPHTWMLPSAHLGIKIDNALIDLASIDPNDGSLDFNEFLFAFV--V 666

601 HKKEQDENIR 612

667 HXLDMKQDPVNR 678

RESULT 13

Q4SHR7\_TETNG PRELIMINARY; PRT; 671 AA.

Q4SHR7; 13-SEP-2005 (TREMBLrel. 31, Created)

DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)

DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)

DT Chromosome 5 SCAFI4581, whole genome shotgun sequence.

DE (Fragment).

GN ORFNames=GSTENG00018033001;

OS Tetradon nigrovittidis (Green puffer).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorphi; Acanthopterygii; Perciformes; Tetraodontiformes;

OC Tetraodontidae; Tetraodontidae; Tetraodon.

NCBI\_TaxId=99883; [1]

RNA NUCLEOTIDE SEQUENCE.

RA Jallou O., Aray J.M., Brunet F., Petit J.L., Strange-Thomann N.,

RA Maucell E., Bounneu L., Fischer C., Oroul-Costaz C., Bernot A.,

RA Nicoud S., Jaffe D., Flesher S., Lutfalla G., Dosset G., Segurens B.,

RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,

RA Anthonard V., Jubin C., Castell V., Katinka M., Vacherie B.,

RA Blemont C., Skalli Z., Catolico L., Poulain J., De Bernardinis V.,

RA Cruaud C., Duprat S., Broctier P., Goutanceau J.P., Gouzy J.,

RA Parra G., Lardier G., Chaplie C., McKernan K.J., McEwan P., Bosak S.,

RA Kellis M., Wolff JN., Guigo R., Zody M.C., Mesirov J.,

RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,

RA Landet V., Schacher V., Quetier F., Saurin W., Scarpelli C.,

RA Winkler P., Lander E.S., Weissenbach J., Roest Crolius H.;

RT "Genome duplication in the teleost fish Tetraodon nigrovittidis reveals

RL the early vertebrate proto-karyotype.";

RL Nature 431:946-957(2004).

RN	[2]
RP	Nucleotide sequence.
RG	Genoscope; Whitehead Institute Centre for Genome Research;
RL	Submitted (FEB-2004) to the EMBL/Genbank/DBJ databases.
CC	-! CAUTION: The sequence shown here is derived from an
CC	EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC	preliminary data.
DR	EMBL; CAAB01014581; CAP99815.1; -! Genomic_DNA.
FT	NON TER
FT	1 671
SQ	SEQUENCE 671 AA; 77077 MW; 5AAE1BF922697ED7 CRC64;
Query Match	36.0%; Score 1238.5; DB 2; Length 671;
Best Local Similarity	37.7%; Pred. No.1.le=66;
Matches 254;	Conservative 122; Mismatches 212; Indels 85; Gaps 7;
OY	5 ATRAPAFQKKYRRHQARREMRRCNMIOFONLATASGDOALYKFFNDILKIMPOAG 64
DB	1 AIKAVALIQRFRRRYMATEMBRBRYYNIPOSIEVASDOLQLSPFSFMINTTLING 60
OY	65 RK---NOYGSAHVSVDKDD--DYVERFDIYNAKIELPRXNHIDLLDVPAKKGN 118
DB	61 SGPDLISGLMDASDSMMENECNFPSVEPTTGPRISFLSLPDMNALSAFERKO-- 118
OY	119 RLHPKYVALILBRAKSLIKQLPNISPVSTAVSOQVTVCGDLHGKLDLLLVYLHKNGLRSS 178
DB	119 TLHARYVLQILLYETRKLLKQMPPNVHLSTVTYKEITTCGDLAGQDLDLLIFYKNGLPSA 178
OY	179 SNPFVFNQDPFDROKRGLEFVLLLLLSLYLAFPNAVPLNRGNHESVMNARGFIREVSK 238
DB	179 ETPEVFNQDFPDROKNSLEVVVLLFAFYLLLPDWMHNRGNHBHLMNLARGFTKEVMOK 238
OY	239 YPRNHKRIATIDEVYRNLPFGSIVNSRVLVHGGPSDSTLDIKSIDRCK----- 290
DB	239 YKHGCEILQPDQVFSLLPVAITIIDGKITLVHGGISDQTDLPUSALSRRKHVSHSEKL 298
OY	291 -----YVSIIRPPL----- 299
DB	299 HFQPIINRKRCKLKRGDSTAANSSSSSSSSSSSSISLGSLKTPPCHTPHRHSSPTLS 358
OY	300 -----TDEGPLDKTEMQOIFDIMSDPOLATMGCVPTNLGAGVFGPDVT 344
DB	359 CSGSAGFQLPFLDSLAPSPPOHEBEMQIVDILMSDPVAPNGCSBPNTFRGGCYFGPDVT 418
OY	345 DNFIQRRLSVITISHHECKPKGHFPMHNKIITPSANYYAIASNGKAYIRLNQGLMPH 404
DB	419 RRLLOQQLILLSHHECKQGEGYEICHSQGVITTFPSANNYEEBSGNRAVYIKVGRELMPR 478
OY	405 FVOYTISAAGOTKRLSFKOMGIVESSALKELAVMMRDHRDELDEFKRYDPEKDSGYIS 464
DB	479 FYQY-QYSRFLTKLTIGRVAAAGSNALAKEKLFTHRSBLLIGFOQYDHNNTOTILVS 537
OY	465 HWCKVMENVTKGLPWRLLDKLAPGDSQKVNTNRITLDLDTVDYLEAEADSGVMDAL 524
DB	538 EMAGVULETGELTELPMWTLRPHLARLASDGVEHYHSCPEDMEPGIPLOVTPNLA--SAL 595
OY	525 YANKASLVAIINIIDAONGSEITLDEFEPTALIDLVAHNPAGYSAEMLEKGRMDLNDG 584
DB	596 FRFKDIIITNIIDKHSGSLISTEFRTWHLSAHGKIDNRPAIDARSIDFNKG 655
OY	585 KVDLNEFLAEARL 597
DB	656 NIDNEFLAEARV 668
RESULT 14	
ID	QARUL1.TEING PRELIMINARY; PRT; 797 AA.
AC	QARUL1
DT	13-SEP-2005 (TREMBLrel. 31, Created)
DT	13-SEP-2005 (TREMBLrel. 31, Last sequence update)
DT	13-SEP-2005 (TREMBLrel. 31, Last annotation update)
DE	Chromosome 12 SCAP14996, whole genome shotgun sequence.

(Fragment).  
 ORFNames=GSTENG0028998001;  
 GN Tetradon nigroviridis (Green puffer).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthopterygii; Acanthopterygii; Percormorpha; Tetraodontiformes;  
 OC Tetraodontidae; Tetraodontidae; Tetraodon.  
 NCBI\_TaxId=99883;  
 [1]  
 NUCLEOTIDE SEQUENCE.  
 RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,  
 RA Maucud S., Bouneau L., Fischer C., Ozouf-Costraz C., Bernot A.,  
 RA Micaud S., Jaffe D., Fisher S., Lutfalla G., Dosat C., Segurens B.,  
 RA Basilia C., Salanoubat M., Levy M., Boudet N., Castellano S.,  
 RA Anthouard V., Ushin C., Castelli V., Katicka M., Vacherie B.,  
 RA Blemont C., Skalli Z., Cattelico L., Poulain J., De Berardinis V.,  
 RA Cnaud C., Duprat S., Broctier P., Coutanceau J.P., Gouy J.,  
 RA Parra G., Lardier G., Chapelle C., McKernan K.J., McGowan P., Bosak S.,  
 RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mestrov J.,  
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,  
 RA Lander V., Schachter V., Queller F., Saurin W., Scarpelli C.,  
 RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.,  
 RT "Genome duplication in the teleost fish Tetradon nigroviridis reveals  
 RT the early vertebrate proto-karyotype.";  
 RL Nature 431:946-957(2004).  
 RN [2]  
 NUCLEOTIDE SEQUENCE.  
 RG Genoscope, Whitehead Institute Centre for Genome Research;  
 RC Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; CAAB01014996; CAG08121.1; -; Genomic\_DNA.  
 FT NON\_TER 1 1  
 FT 797 797  
 SQ SEQUENCE 797 AA; 91973 MW; 89F01F058C820730 CRC64;

Query Match 35.8%; Score 1232; DB 2; Length 797;  
 Best Local Similarity 32.5%; Pred. No. 3.5e-66;  
 Matches 261; Conservative 135; Mismatches 190; Indels 218; Gaps 13;

5 AIRAATFIOKMYRHOAREMORCMQVQNTLEVASSEDOAEVYFNDLIGKMPQAG 64  
 1 ALAAATVIGQMYQYARSKMRKCTWHITQSTIEYSEQAIKLSFNLMTMNPSSN 60  
 65 RKNQYOGSAHVSLDD--KODLVEEFQDIYNAKI-----ELPIKKNHIDLIDVPR 113  
 61 ERRLI---SHIFRENEVCRAEMERYFCYNIIEVPEIYSGPHLTFPLTVGQNGVIAFR 117  
 114 KRGKGNLHPKYVALILBEAKSKIKQLPNISPVSTAVSQVTV-----156  
 118 NKR--QLHSRYIIQLLETWKLRLMPLNIRISTCHSKETITCGELQEGFGNGLCREIV 175  
 157 -----GDHGLKLDLILVLAHK-----172  
 176 GFTLPPTVLTWVAAPIGDHLGQLEDDLIFYVRNPLPLFYMFSPTVVFLHVTYQTSN 235  
 173 -----NGLPSSNPYVFGDFVDRGKRGLEVVLLLSLYLAFPAVFLNR 217  
 236 NNNKNRNVFNCAALQNGMPLEKPYFNGDFVDRGDSITIIIVILAFPLVYSQVYANR 295  
 218 GNEHDSVNAARYPFIREVSKYPRNNKRIIAAFIDEVTRMLPLGSLNRYLVHGGPSDS 277  
 296 GNEHEDIVNLRYPFTKTEVLTKYHGGKRIKLQKIFSWPLATVLDQKVLVHGGISDL 355  
 278 TSLDLKSIDRGKYVSLRPP-----LITGSEPL 305  
 356 TDLGLAKLDRNHYVSLRPPKRGYVSACTISDSVDDEVCSNKKI PQRTSLTYARPL 415  
 306 D-----K 307  
 416 GRTTSFQNRSLQDFSDRIRLPAENELRLRRRREPIFILSSEKTIINPAASDSVKSSELK 475

308 TEMQOIFDMMSPQATMGCVPTLIRGAGVMPGPDVTDNPIORHRLSYVIRSHCKPNHG 367  
 476 DEKQILDLIMSDPMIQDGCIPNEVRGGCTWOPDYTEETLNKNIQLIIRSHCKQBGY 535  
 368 EFMHDKII-----TFPSANYYAIGSNKAYIRLNOL 401  
 536 EFCNKNKVIQGLPGPGPRVLPGHQLKTVPLQVLLTFPSANNYYDVGSNRGAVVLGPNL 595  
 402 MPEFVQYISAAGTKRLSKRMGIVSSALRELAVRMGRDHEDEBRKIDPKDSGYI 461  
 596 VPVIOY-QANSTRELISMGOSIGRTERSALKVIREQLPAHKSDLISAFQPTENTYGLV 654  
 462 SISHWKWKNENTVTKGLPRLIRDKLAPGTDSQ-KYNVRRITLDLITDVLLEADGMSV 520  
 655 SLCDMAASAVSWHGLPWRMLSSQSLTTSKSGDMSTYHWFDEL--AIKGNIDGLQ 711  
 521 -----MDALYANKASVIAFNIIDANSGEITLDEPFIADLLVAHMPGASVKAEMLEKC 575  
 712 TLYRHLFTLYRHSRTLETIFRIIDKNSGFIEMEDPRKTKLISLYLKNHISDALSDLA 771  
 576 RMDLNGDKVLDNFELEAFRLSD 599  
 772 VTIDSQDSIDIDFEBAFRLTD 795

## RESULT 15

ID 08MYR2\_CABEL PRELIMINARY; PRT; 572 AA.  
 AC 08MYR2;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DT Phosphatase with ef hande protein 1, isoform b.  
 GN Name=pef-1; ORFNames=F23H11.8;  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 NCBI\_TaxId=6239;  
 [1]  
 NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=Br1etol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RG The C. elegans sequencing consortium;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology.";  
 RL Science 282:2012-2018 (1998).  
 DR EMBL; AF003389; AAM22065.1; -; Genomic\_DNA.  
 DR HSSP; P36873; 1IT6.  
 DR Ensemble; F23H11.8; Caenorhabditis elegans.  
 DR Wormbase; WBGene0003969; pef-1.  
 DR WormPep; F23H11.8b; CR30662.  
 DR GO; GO:0005509; F:calcium ion binding; IEA.  
 DR GO; GO:0016787; F:hydrolase activity; IEA.  
 DR InterPro; IPR011992; EF-Hand type.  
 DR InterPro; IPR002048; EF\_Hand\_Ca\_bd.  
 DR InterPro; IPR004843; W-pesterase.  
 DR Pfam; PF00036; ehand\_2.  
 DR Pfam; PF00149; Metallophos; 1.  
 DR PRINTS; PR00114; STRPHRTASE.  
 DR SMART; SM00054; EFh; 3.  
 DR SMART; SM00156; PP2Ac; 1.  
 DR PROSITE; PS00018; EF\_HAND; UNKNOWN 2.  
 DR PROSITE; PS00125; SER\_THR\_PHOSPHATASE; UNKNOWN 1.  
 KW Complete proteome.

SO SEQUENCE 572 AA; 65150 MW; 504039A0685C98CD CRC64;

Query Match 33.3%; Score 1147; DB 2; Length 572;  
 Best Local Similarity 41.0%; Pred. No. 3.3e-61;  
 Matches 227; Conservative 106; Mismatches 174; Indels 46; Gaps 10;

81 KODLVEEFQDIYNAKI-----LPIRKNHIDLIDVPRKKNRRLH 121

```

Db      27 KPSLMSBGEIVYKMLEBDSPTNVIDIDRNVYKGPITSLPLDKPQVAKMEAF--KVNKVLH 84
Qy      122 PKYVALLIEBAKSKLKQLPNISPVSTAVSQOYTVCGDLHGKDDLLVYTHKGLPSSNP 181
Db      85 PKYVALLIEBAKSKLKQLPNISPVSTAVSQOYTVCGDLHGKDDLLVYTHKGLPSSNP 144
Qy      182 YFNGDFVDRGKGLKGLVLLLSLYLAFPAVAFNAGNHEDSVNNAARYGFIREVSQYR 241
Db      145 YIFNGDFVDRGKGLKGLVLLLSLYLAFPAVAFNAGNHEDSVNNAARYGFIREVSQYR 204
Qy      242 NHRKILAFIDEVYRWLPDGSVINSRLVYHGFSDSLDLKSIDRGKYSILRPPLTD 301
Db      205 LSTPIRLLEDFVSMPIPIATIIDRIDFVYHGSGISDQTEVSKLDKIPRHFQSVLRPYNK 264
Qy      302 GSEPLD-----TEWQOIFDIMMSDPQATMGCVPTTLGAGVWEPDVTDFLQHRHS 354
Db      265 GMSSEKENSANVNDKQMLDIMMSDPKQKQKQWPFVFGGSGYFGADITASFLHGGR 324
Qy      355 YVIRSHCKPNGHGFPMNDKILITIFASNYVYIAGNKAYIR-LNNQLMPHFVOYISAAS 413
Db      325 LVIRSHCKPNGHGFPMNDKILITIFASNYVYIAGNKAYIR-LNNQLMPHFVOYISAAS 382
Qy      414 QTRKLS-FKQRMQIVSASAKELAVMRDHRDELDEFRKYDPKQSGYISISHCKWMEN 472
Db      383 KTHRKSTLBERLGVVRESAVKELKELSGFHTDLQKEFEIMDIKRSKGLPILKMSDQYER 442
Qy      473 VTKGLPMLLDKLP-GLDSQKQVYNTLIDLDTVYLEABADQMSWMDLYANKSL 531
Db      443 ITGLNLPMTALAPKQATLSEBDGCVYKEDRRJAQVGTHAQEKD---YVESLYRKSTL 499
Qy      532 VAFNFIIDNDSGEITLDFETAILDLVAHMPGAYSKAM-----LEKRMMDLNGDGKY 586
Db      500 ETILFRMDKDNQSGVSMKEPIADACEVL-----GKTKRPLQDTYISQIAESIDPNQGT 554
Qy      587 DINEFLAEFRSLD 599
Db      555 DINEFLAEFRSLD 567

RESULT 16
Q4R441_MACFA PRELIMINARY; PRT; 405 AA.
ID Q4R441_MACFA PRELIMINARY; PRT; 405 AA.
AC Q4R441;
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Testis cDNA clone: Q6A-12561, similar to human protein phosphatase,
DE EF hand calcium-binding domain 1 (PP6F1), transcript variant 1.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopithecoidea; Cercopithecinae; Macaca.
NCBI_TaxID=9541;
OX NCB1_TaxID=9541;
RN NM_001001001.1
RP NUCLEOTIDE SEQUENCE.
RA International consortium for macaque cDNA sequencing, analysis;
RT "DNA sequences of macaque genes expressed in brain or testis and its
RT evolutionary implications."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
[2]
NCUBOTIDE SEQUENCE.
RA Oseada N., Hirata M., Tanuma R., Kusuda J., Hida M., Suzuki Y.,
RA Sugano S., Gojobori T., Shen J.C.-K., Wu C.I., Hashimoto K.;
RA "Substitution rate and structural divergence of 5'UTR evolution:
RT Comparative analysis between human and cynomolgus monkey cDNAs."
RT Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB179074; BAE02125.1; mRNA.
SQ SEQUENCE 405 AA; 46636 MW; C139899F20B234B CRG64;

```

Query Match 24.4%; Score 839; DB 2; Length 405;  
 Best Local Similarity 41.3%; Pred. No. 1e-42;  
 Matches 159; Conservative 79; Mismatches 111; Indels 36; Gaps 4;

```

Qy      3 ENAIRAIFIOKMYRHRQARRBMRQCMQIFONLEVAESODQAEIYKFPNDLIK----- 57
Db      14 ETSLEPAALIQWNYGYRKLAKQRYALTTFQSLIEYDQSQMLSTFFSFWFNAYAHV 73
Qy      58 HMPQAGRKNOYQSAHVSVDKXDLVEEFGDIYNATIEPIRNKHIDLIDVRKQKG 117
Db      74 HNGEIELNQSLESQDLRDRIHDVYDLIDVPSYNGRPLQFPLCTCTDIDLLBAKQO- 132
Qy      118 NRIHKRYALLIRBAKSKLKQLPNISPVSTAVSQOYTVCGDLHGKDDLLVYTHKGLPSSNP 177
Db      133 -ILAHVYLEVLFEYKYLKQMPNPTVQTSKELTICGLDHGLDLDLFLFYNGGLPS 191
Qy      178 SGNPVPFNGDFVDRGKGLKGLVLLLSLYLAFPAVAFNAGNHEDSVNNAARYGFIREVS 237
Db      192 SGNPVPFNGDFVDRGKGLKGLVLLLSLYLAFPAVAFNAGNHEDSVNNAARYGFIREVS 251
Qy      238 KYPRNKRILAFIDEVYRWLPDGSVINSRLVYHGFSDSLDLKSIDRGKYSILRP 297
Db      252 KYKLGKRIILQLEBELVAMLPITGITVDNEILVYHOGISFTDNLNHRVERKMSVLIP 311
Qy      298 PL-----TDGEP--LDKTEWQOIFDIMMSDPQATMGCV 328
Db      312 PLETNRDNDTDSKNNKGVTFNAHRIRKTNQSPSEHLFKHEWQIIDLMSDPKQKQCF 371
Qy      329 PNTLRAGVWFPDVTDFLQHRHL 353
Db      372 PNTGRGCGCFQDPVTSKILNKYQL 396

RESULT 17
Q68EP0_XENTR PRELIMINARY; PRT; 493 AA.
ID Q68EP0_XENTR PRELIMINARY; PRT; 493 AA.
AC Q68EP0;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DR PP5C-prov protein.
CN Name=pp5c-prov;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus; Silurana.
NCBI_TaxID=8364;
OX NCB1_TaxID=8364;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUS=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altchul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Meng J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huylk S.W.,
RA Villalon D.K., Mizny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting R.W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield J.S.N., Krzywinski M.I., Skalska U., Smalins D.B.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
NCUBOTIDE SEQUENCE.
RA TISSUS=Embryo;
RC TISSUS=Embryo;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
CC -I- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +

```



[illegible]

CC Xenopodinae; Xenopus; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TTSBUSE-Ovary;  
 CC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feilgold R.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldi M.P., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.U., Ussidi T.B., Toshiyuki S., Carrinci P., Prange C.,  
 RA Rata S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Boak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Mulny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton S., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield V.S.N., Krzywicki M.T., Skalska U., Smalhe U.E.,  
 RA Schmeich A., Schein J.E., Jones S.J.W., Marra M.A.,  
 RA "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TTSBUSE-Ovary;  
 CC MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.,  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT Initiative";  
 RL Dev. Dyn. 225:384-391(2002).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TTSBUSE-Ovary;  
 RA Klein S., Strausberg R.;  
 RL submitted (JUN-2004) to the EMBL/Genbank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +  
 phosphate.  
 CC -1- SIMILARITY: Belongs to the PPP phosphatase family.  
 CC EMBL; BC073033; AAH73033.1; -; mRNA.  
 CC SMR; Q6GPS6; 20-493.  
 DR GO; GO005737; Cytoplasm; IEA.  
 DR GO; GO005634; Cnucleus; IEA.  
 DR GO; GO0016787; Phosphatase activity; IEA.  
 DR GO; GO0004721; Phosphoprotein phosphatase activity; IEA.  
 DR GO; GO0006470; Protein amino acid dephosphorylation; IEA.  
 DR InterPro; IPR004843; M-esterase.  
 DR InterPro; IPR011236; PPTase\_5.  
 DR InterPro; IPR006186; T\_Phtase\_aph.  
 DR InterPro; IPR001440; TPR.  
 DR InterPro; IPR011990; TPR-like\_helical.  
 DR Pfam; PF00149; Metallophos; 1.  
 DR Pfam; PF00515; TPR\_1; 3.  
 DR PIRSF; PIRSF033096; PPTase\_5; 1.  
 DR PRINTS; PR00114; STPHPTASE.  
 DR ProDom; PD000252; T\_phtase\_aph; 1.  
 DR SMART; SM00156; PP2AC; 1.  
 DR SMART; SM00028; TPR; 3.  
 DR PROSITE; PS00125; SER\_THR\_PHOSPHATASE; UNKNOWN\_1.  
 DR PROSITE; PS50005; TPR; 3.  
 DR PROSITE; PS50293; TPR\_REGION; 1.  
 KW Hydrolase; Iron; Repeat; TPR repeat.  
 SQ SEQUENCE 493 AA; 56276 MW; 127DC9409402C297 CRC64;

```

Db      182 VDMFELMGPYKQKGLHRCICLVQMLVQVNDILSKPLSVLEVLSKSCQVTVCCGDTHGCF 241
      164 DDLVLVHLKNGLPSSSNPVYFNGDFVDRGKGLLEVLILLSLYAFPNVAVLNRGNHDS 223
      242 YDLNNTFHLNGLSENNPYIFNGDFVDRGSPSVSEVITLTFGLYPAHFLHNRGNHETD 301
      224 VNNARGLFREVSKYPRNRKRIILAFIDERYRMPLGSLVNSVTLVHGCF--SDSTSLD 261
      302 TNNQMGFGESEVAKT---SAQMFQLFSVQMLPLMVCVNRVLNHWGLFSEIDVTLTD 358
      282 LKSIDRGKYSILRPDLTGEPLDKTEWQOIFDIMSDFQATMGCVPTLRAGAVWFGP 341
      359 QINIERNR-----QRP--DSGPM-----CDLMSDPQPDG-RSTSKRGVSCQFGR 402
      342 DVTDFNLFQRRHLSYVTRSHCEKPGNGHEFMNDKTIITFSASNYVALGSKNGAYIRLN-NQ 400
      403 DVTHQFLSENGLDYIIRSHVEKDEGEVSHNGLCVTVFSAFNPYCDQGNKGAYIHLSGSD 462
      401 LMFHPVOY 408
      463 LKRFHQF 470

Db
      463 LKRFHQF 470

RESULT 21
ID Q7QB71_ANOGA PRELIMINARY; PRT; 474 AA.
AC Q7QB71;
DT 01-MAR-2004 (TEMBLrel. 26, Created)
DT 01-MAR-2004 (TEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE ENSANGP0000011234 (Fragment).
GN ORFNames=ENSANG00000008745;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anophelae.
OX NCBI_TaxID=180454;
RN RN
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation.";
RN Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.
RN Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -1- SIMILARITY: Belongs to the PPP phosphatase family.
DR EMBL; AAA0100880; EAA08659.2; -; Genomic_DNA.
DR GO; GO:0005737; C:Cytoplasm; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004720; F:protein amino acid dephosphorylation; IEA.
DR GO; GO:0006701; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR004843; M-pesterase.
DR InterPro; IPR011236; PPTase_5.
DR InterPro; IPR006186; PPTase_5.
DR InterPro; IPR001440; TPR.
DR Pfam; PF00149; Metallophos; 1.
DR Pfam; PF00515; TPR_1; 2.
DR PIRSF; PIRSF03096; PPTase_5; 1.
DR PRINTS; PR00114; STRPHPTABE.
DR ProDom; PD000252; T_Phtase_apah; 1.
DR SMART; SM00156; PP2Ac; 1.
DR SMART; SM00028; TPR; 2.
DR PROSITE; PSS0005; TPR; 2.

```

```

DR PROSITE; PSS0293; TPR_REGION; 1.
KW Hydrolase; Iron; TPR Repeat.
FT NON TER. 1
FT NON TER. 1
SQ SEQUENCE 474 AA; 54287 MW; A7BDD4FD5BCAC17A CRC64;

Query Match 17.9%; Score 614.5; DB 2; Length 474;
Best Local Similarity 36.4%; Pred. No. 5.5e-29;
Matches 150; Conservative 64; Mismatches 155; Indels 43; Gaps 12;

      11 FLOKWRHQRAREMRQRCNMQIFQNLBYA-----SEODAELEYKFENDLIKMPQAAGR 65
      74 YTKAYYRBAALWALGR--FKALADLEFPAKRCFAKXADQYTBCKKVNKILAEKAI 131
      66 KNOYGSANVSLVD-----KQDLVEBFQDLYNAKIELPIRKNHIDLIDVFRKGRNR 119
      132 SVQHQKSVKMKCRDMEFATIEDDYGP--KLENGVTTLTFMKS-----LLEWYKKQ--NK 163
      120 LHKYVALILREAKSLKQLPNTSPVSTAVSQQVTCGDLHGKLDLVLVHLKNGLPSSS 179
      184 LHKFAYRILCDMETLLKTPQPSLVEIPVDEQKFTVCGDTHGQFYDLMLFEINGLPSPT 243
      180 NPVYFNGDFVDRGKRGLEVLILLSLYLAFPNAVFLNRGNHDSVNNARGYFREVESKY 239
      244 NPYLFRNGDFVDRGSPVECEFTLFGFKLYPNHFLAKGNHESFNNQLYGFTGEVYVSKY 303
      240 PRNHRKRIILAFIDERYRMPLGSLVNSRVILVHGCF--SDSTSLDLKSIDRGKYSILRP 297
      304 SQWMAFMFTL---VYNNMLPLCHLIRKKVLYMHGGLFSPKQVNSLDDRSIDRN-----CQP 355
      298 PLTGDGPRDKTEWQOIFDIMSDFQATMGCVPTLRAGAVWFGPDDVTDFNLFQRRHLSYVI 357
      356 P-----BEGMLCELMSDPHPLPGRVPSK--GQVGIQFQPDVTEALKNNMLDYII 404
      358 RSHCEKPGNGHEFMNDKTIITFSASNYVALGSKNGAYIRLN-NQQLMFHPVOY 408
      405 RSHVEVABGEYVDHNGKCTVFSAPNYCDQMKNLGAFITLKGNDLIPKFTTY 456

Db
      405 RSHVEVABGEYVDHNGKCTVFSAPNYCDQMKNLGAFITLKGNDLIPKFTTY 456

RESULT 22
Q4RPT2_TETNG
ID Q4RPT2_TETNG PRELIMINARY; PRT; 874 AA.
AC Q4RPT2;
DT 13-SEP-2005 (TEMBLrel. 31, Created)
DT 13-SEP-2005 (TEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TEMBLrel. 31, Last annotation update)
DE Chromosome 16 SCAR15113, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00035185001;
OS Tetraodon nigroviridis (Green puffer).
OC Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN RN
RP NUCLEOTIDE SEQUENCE.
DR Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
DR Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
DR Michard S., Jaffe D., Fisher S., Lutfalla G., Doesat C., Segurens B.,
DR Desliva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
DR Anthouard V., Jubin C., Castelil V., Katinka M., Vacherie B.,
DR Blemont C., Skalli Z., Catolico L., Poulain J., De Bernardis V.,
DR Cruaud C., Duprat S., Bottier P., Coutanceau J.P., Gouzy J.,
DR Paria G., Lardier G., Chapelle C., McKernan K.J., McMan P., Bosak S.,
DR Kelle M., Wolff J.N., Guigo R., Zody M.C., Mestrov J.,
DR Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
DR Lauder V., Schachter V., Queller F., Saurin W., Scarpelli C.,
DR Wincker P., Lander E.S., Weissbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RN the early vertebrate proto-karyotype.";
RN Nature 431:946-957(2004).
RN [2]

```





Db 653 DKVLGKVSFSEFEQVLTAKOLNINLSABQVRIIM-----RMINSNLSCKKRVQGENDKIDVA 708  
QY 590 EFLBAR-----LSDLRK--EQDENIRRS 614  
Db 709 EFIGKRVCTRLAINKYNNEKVOKLIEIGKHLADSEIANFHYRFEYEQDESENAD 768  
QY 615 TGRSVAKT 623  
Db 769 RKSSVYKS 777

RESULT 24  
081728\_TRYCR  
ID 081728\_TRYCR PRELIMINARY; PRT; 923 AA.  
AC 081728;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE TCC31.31.  
GN Name=TCC31.31;  
OS Trypanosoma cruzi.  
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma;  
OC Schizotrypanum.  
OX NCBI\_TaxId=5693;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=Brener;  
RX MEDLINE=98391765; PubMed=9724326;  
RA Anderson B., Aelund L., Tammi M., Tran A.N., Hohelsel J.D.,  
RA Pettersson U.;  
RT "Complete sequence of a 93.4-kb contig from chromosome 3 of  
RT Trypanosoma cruzi containing a strand-switch region.";   
RL Genome Res. 8:809-816(1998).  
[2]  
RN NUCLEOTIDE SEQUENCE.  
RC STRAIN=Brener;  
RA Anderson B., Boncompagni E.J.;  
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; A0137988; AAN78342.1; -; Genomic\_DNA.  
DR HSSP; P36873; 1176.  
DR GO; GO:0016787; P:hydrolase activity; IEA.  
DR InterPro; IPR011992; BP-Hand type.  
DR InterPro; IPR002048; EF\_Hand\_Ca\_Bd.  
DR InterPro; IPR004843; W-pesterase.  
DR InterPro; IPR006186; T-pesterase\_apah.  
DR Pfam; PF00149; Metallophos; 1.  
DR PRINTS; PR00114; STPHPTASB.  
DR ProDom; PD000252; T\_phlase\_apah; 1.  
DR SMART; SM00156; P2AC; 1.  
DR PROSITE; PS00125; SER\_THR\_PHOSPHATASE; UNKNOWN\_1.  
SQ SEQUENCE 923 AA; 104798 MW; 95B6646B32B2143 CRC64;

Query Match 17.3%; Score 594.5; DB 2; Length 923;  
Best Local Similarity 33.0%; Pred. No.2,1e-27;  
Matches 174; Conservative 73; Mismatches 206; Indels 75; Gaps 15;

QY 128 IIREAKSLKQLEPNI-----SPVSTAVS-----QQVTCGDLHGKLDLVLVLRNGLP 176  
Db 299 VFIDAMSHLNTMNVNVRISLPFGARVNGRINGSKVVVGDHGLADLHLILKCGMP 358  
QY 177 SSSNPVFNQGFVDKRGKGLVLLLSLYLAPPAVFLNRGNHDSVNNARYGFTREYE 236  
Db 359 NEGTYYIFNGDFDRGANGVEVLLILFSLMLACPKYVTLNRGNHCDYNNMDRYGPFVEVS 418  
QY 237 SKYPRNKRILAIIDEYVRLPLGSVLSRVLLVHGGF--SDSTSLDLKSIDRGKYSI 294  
Db 419 TKIDRN---VRLVQRCFLPLATITIGKVPVYHGGELPRKGVNTEDISRIQRPROI-- 473  
QY 295 LRPEPLDGEPLDRTKTEWQOIF--DIMSDPQATMGCVPTNLGAGVWFGPDVDTNPLQRHRL 353  
Db 474 ---FMED---YSQPEEDBEIRFODLLMSDPVEDLQGMRSRPGAGVVGADVTQFLOQNGL 527  
QY 354 SYVIRSHCEKCPNGHEFPHNDKITTIPSASNYVAIGSKGAV--IRLNQMLPHFVQYISAA 412

Db 528 ELVIRSHCEKAGYEENHDKGLITVPSASNYDGPETNPFSGPAVFGVDNBPESHYQVAE 587  
QY 413 SQT-----RLSPKQRMGIVESA-----LKLAVMRDRHDELD 448  
Db 588 DVEVQTLVDLGTFTPTLGRVSSFATLSQSKLLRRARADVLRRLRIYGRRLRLA 647  
QY 449 EPRKIDPKDSGYTISISHCKVWENVTYKGLPRLRLDKLAPGDSQKATNNTL----- 502  
Db 648 YPAKLDRTKSGSVWKLEWYEAARMVLTLDLPWFPLEGYLVADENRTVWYSHFLVXPHNE 707  
QY 503 --DLDTVILKEADGMSVMDALYANKASIVAFINIIDANSGETLTLEFEPAIDLLVA 560  
Db 708 FOPLMINDWIOQH---CHHLTQOQRANHRQY---VAAFPNGQVSNPECSVIRAIIDY 760  
QY 561 HMFAGSKAEMLEKCRMDLNGDKYDNEF-----LEAFRLSLDR 602  
Db 761 TM-----SEAQLQLFYLDGGTGHTDGRKPVNMLEMAAYPLSDTLR 804

RESULT 25  
0512P5\_ENTHI  
ID 0512P5\_ENTHI PRELIMINARY; PRT; 473 AA.  
AC 0512P5;  
DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DE Protein phosphatase, putative.  
GN ORFName=90.t00028;  
OS Entamoeba histolytica HM-1:IMSS.  
OC Eukaryota; Entamoebidae; Entamoeba.  
OX NCBI\_TaxId=294381;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=HM-1:IMSS;  
RX PubMed=15729342; DOI=10.1038/nature03291;  
RA Loftus B., Anderson I., Davies R., Altmak U.C., Samuelson J.,  
RA Amedeo P., Roncaglia P., Berriman M., Hitt R.P., Mann B.J., Nozaki T.,  
RA Suh B., Pop M., Duchene M., Ackers J., Tannich E., Leipe M.,  
RA Hofer M., Bruchhaus I., Willhoeft U., Bhattacharya A.,  
RA Chillingworth T., Churcher C., Hance Z., Harris B., Harris D.,  
RA Jagels K., Moule S., Mungall K., Ormond D., Squares R., Whitehead S.,  
RA Quail M.A., Rabinowitsch E., Norbertczak H., Price C., Wang Z.,  
RA Guillen N., Gluchrist C., Strop S.E., Bhattacharya S., Iohia A.,  
RA Foster P.G., Sichteritz-Ponten T., Weber C., Singh U., Mukherjee C.,  
RA El-Sayed N.M., Petri W.A., Clark C.G., Embley T.M., Barrett B.,  
RA Fraser C.M., Hall N.;  
RT "The genome of the protist parasite Entamoeba histolytica.";  
RL Nature 433:865-868(2005).  
CC -1- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL; AAPB01000329; BALT4840.1; -; Genomic DNA.  
SQ SEQUENCE 473 AA; 53606 MW; 665B460BC0ED93A8 CRC64;

Query Match 16.9%; Score 582.5; DB 2; Length 473;  
Best Local Similarity 35.4%; Pred. No.4,8e-27;  
Matches 143; Conservative 74; Mismatches 134; Indels 53; Gaps 14;

QY 11 FLOKTYRRHQ-----REMQRCWQIFQNLIVSBDQALYFPPNDLIGHMQAAG 64  
Db 68 FVKGYYRQASAPAAQLQLEAISAC-----ERAKKLSPPD-----GMISMKGLREK-- 115  
QY 65 RKNQYQSAHVSGLDDKDL-----VEFGDIYNAKIELPRKNHLDLIDVPRKGN-- 118  
Db 116 RRRQLPFEA-ISEVBEKQRTISMRDIDASTSTIKIBDKITTEENYQ---EIVAMSNQ 171  
QY 119 ---RLHPRYVALIIREAKSLKQLEPNI-----SPVSTAVSQQVTCGDLHGKLDLVLVLRNGL 174  
Db 172 YGKCAIHLKYCRLIIEBAGIILKKRVNLEIINTQ--GKRWITVGDHIGCFDHLHPEING 230  
QY 175 LPSSNPVFNQGFVDKRGKGLVLLLSLYLAPPAVFLNRGNHDSVNNARYGFTREYE 234

Db 231 LPSEDNVYLLENQDFVDRGSGVGEVCLTFSPHIIYENSVPLARNGHETRAMNANYGEFE 230  
 Qy 235 VESKTPRNKRLLIAFLDEYTRMLPLASVLANSVLLVHGQFSDS-TSLDLKSIDRGKTVS 233  
 Db 291 VKTKY--NENVYLAFFSD-VFMQLPYCHVIDSKVFLVHGQIDPMYSLDLIDIKIKRG---- 343  
 Qy 294 ILAPPLDSEPLDKTEWQQLFDLMGSDPAAITGCPNTLIRAGGVFSGPDYTDNPLQNHRL 353  
 Db 344 -----CPQGSIASALALVMDPPTSGSSGR-SIRQCGKSGFGPDITHPLDGNLT 391  
 Qy 354 SYVRSHECKPENGHEFMDNKILITTSASNYVAIGSNQAVYRL 397  
 Db 392 QYIVRSHENKNGIEMGARGRLITVSAANYCDQNNNGAIIYHV 435

```

1  RESULT 26
2  Q9NES8_CABEL
3  ID Q9NES8_CABEL PRELIMINARY; PRT; 496 AA.
4  Q9NES8;
5  DT 01-OCT-2000 (TREMBLrel. 15, Created)
6  DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
7  DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
8  DE Hypothetical protein Y39B6A.2.
9  GN ORFNames-Y39B6A.2;
10 OS Caenorhabditis elegans.
11 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
12 CC Rhabditidae; Pelodierinae; Caenorhabditis.
13 NX NCBI_Taxid=6239;
14 [1]
15 RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
16 RP STRAIN=Bristol N2;
17 RX MEDLINE=99069613; PubMed=9851916;
18 RG The C. elegans sequencing consortium;
19 RT "Genome sequence of the nematode C. elegans: a platform for
20 investigating biology.";
21 RL Science 282.2012-2018 (1998).
22 DR EMBL, AL332948; CA51076.2; -; Genomic_DNA.
23 DR PIR, T45058; T45058.
24 DR HSSP; P53041; 1A17.
25 DR InFACt; Q9NES8; -.
26 DR Ensemble1; Y39B6A.2; Caenorhabditis elegans.
27 DR WormBase; MGene00012655; Y39B6A.2.
28 DR Wormpep; Y39B6A.2; C36619.
29 DR GO; GO:0005737; C:cytoplasm; IEA.
30 DR GO; GO:0005634; C:nucleus; IEA.
31 DR GO; GO:0005488; F:binding; IEA.
32 DR GO; GO:0016787; F:hydrolase activity; IEA.
33 DR GO; GO:0004721; F:phosphoprotein phosphatase activity; IEA.
34 DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
35 DR InterPro; IPR004843; M-peptidase.
36 DR InterPro; IPR011235; PPTase_5.
37 DR InterPro; IPR006186; T_phatase_apah.
38 DR InterPro; IPR001440; TPR.
39 DR InterPro; IPR011990; TPR-like_helical.
40 DR Pfam; PF00149; Metallophos; 1.
41 DR Pfam; PF00515; TPR_1; 2.
42 DR PRINTS; PRSFR03096; PPTase_5; 1.
43 DR PRINTS; ER00114; STEPHITASE.
44 DR PRODom; PD000252; T_phatase_apah; 1.
45 DR SMART; SM00028; TPR; 2.
46 DR PROSITE; PSS0005; TPR; 2.
47 DR PROSITE; PSS0293; TPR_REGION; 1.
48 KM Complete proteome; Hypothetical protein.
49 SQ SEQUENCE 496 AA; 56462 MW; 2C872FF1700384EB CRC64;
50
51 Query Match 16.9%; Score 581; DB 2; Length 496;
52 Best Local Similarity 33.0%; Pred. No. 6.3e-27;
53 Matches 137; Conservative 77; Mismatches 167; Indels 34; Gaps 12
54
55 3 ENAIRAI---FQKKYRRHQARRRQK-RCKMQIFQNLRTASBDQAEIYKF--FNDLI 56
56 DMI--AIDSYVYKGYGRATANMAGRTFKKALTIDYQAVVKKCPNDKDAKPFECSEKIV 143

```

```

0Y 57 KHMPOAGKQKQYQGSANVSVLDDDDDLVBERGFIYNAKIELPIRKHNDILLDIPPKKR 116
Db 144 ROKCFEALISTDHDKXTVAETIDINAMAIKSYD-GRLEBDKTYKEFVLIQILKTFENQQ 201
0Y 117 GNRLEPKYVALIIRERAAKSLKOLPNIISPVSPAIVQOYVCGDLHGKDDLLVYLHKNGLP 176
Db 202 --KLMKCAFYKULLEFYNYVLSLPIFWBETIYPTQKKTPICODVHQYYDLCNIFZINGYR 259
0Y 177 SSBNPVFNQDFVDRGKGLLEVLLLSLYLAFNAAVFLNRGNHDSVMNARYGFIREVE 236
Db 260 SETNPFLFNGDFVDGSGFSVEITFMIIGFKLLPHNFPMGSGNBSVMNKNYGFEBEYK 319
0Y 237 SKYPRNKHKILAFIDBYRYRKLPLGSLVNSRVLVYHGF--SDTSISDLISIDRGKTV81 294
Db 320 AKT--OQMCMPTETFCMLPLCHLINBKLFVCHGGLFKBDGVTLLDIRKTDRN----- 371
0Y 295 LRPRLDGRPLDKTEMQOIFDIIMSSDPAITMGCVPNTLRGAGVAFGVDVTDNPLQHRRLS 354
Db 372 -RPPDEG-----IMCDLLMSDPPINGRSPSK-RGGGCGFGADVTSKCEITNGIE 420
0Y 355 YVIRSHCKPKNHGFPHNDNKIITTPSAANYAISNKAQVYRL-NNOLMHPHFOY 408
Db 421 YVIRSHCKPKNHGFPHNDNKIITTPSAANYAISNKAQVYRL-NNOLMHPHFOY 475

```

Query Match	Best Local Similarity	Score 578;	DB 2;	Length 829;
27.4%	Pred.No. 1.9e-26;			

DR InterPro; IPR004843; M-gesterase.  
 DR InterPro; IPR011236; ppIase\_5.  
 DR InterPro; IPR001440; TPR.  
 DR InterPro; IPR006186; T\_phcase\_apah.  
 DR Pfam; PF00149; Metallophos; 1.  
 DR Pfam; PF00515; TPR\_1, 3.  
 DR PIRSF; PIRSF03096; ppIase\_5; 1.  
 DR PRINTS; PR00114; STPHPTASH.  
 DR SMART; SM00156; PP2ac; 1.  
 DR SMART; SM00028; TPR; 3.  
 DR PROSITE; PS5005; TPR; 3.  
 DR PROSITE; PS50293; TPR\_REGION; 2.  
 DR SEQUENCE 594 AA; 69196 MW; 63E3E3750664181C CRC64;  
 SQ  
 Query Match 16.8%; Score 577.5; DB 2; Length 594;  
 Best Local Similarity 35.9%; Pred. No. 1,36-26;  
 Matches 151; Conservative 77; Mismatches 144; Indels 49; Gaps 17;  
 QY 6 IRAAIFQKMYRRHQAREMRQRCNMQIFQNLBYASE--QDAELYKFNDLIKMPQAA 63  
 DB 184 IDEAIKINPYX---AKAYYRKGSYLLSLDKRASECFOKVLKLTQDKNSBLK-LKQCK 238  
 QY 64 GR--KNOYQSAHYSLVDDK---DDLVBEPGDIIVAKIELPT-RKQHIDLLDVFKKR 116  
 DB 239 KLIFEQGFQKAIK--LEQKMPYETVLVLSLKEN--MEAPYIDRNINLV--DFLKVA 291  
 QY 117 -----GNRLHPKVVALILREAAKSLKQLNPISPVSTAVASQOYVTCGLDGKLDLLVL 170  
 DB 292 DYIETPNKKLKKCVCAIVLDVYLKLKSLPTLVYLNLEBDETLTICGVHGOYIDLNTIM 351  
 QY 171 HKNGLPSSSNPYENVGDFVDRGKGLFVLLLSLYLAFPNVAFNLGNHDSVMNARYG 230  
 DB 352 KINGYPEBKNSYLVNGDFVDRGSPSEVVIIFLYLAKLTFNNVYLTGNGHETDNNTIYG 411  
 QY 231 FIRVESKTPRYNPKRIIAFIDVRYRWLPGLSVLNSRVLYHGGF--SDSTSLDLIKSIDR 288  
 DB 412 FLGLGLOEKY--DEKTHVLFSDS-FKPLPLAYLVANKNIFICHGGIPSKITDTTLBIEKIDR 468  
 QY 289 GKVYSILRPLTDSERPLDKTEWQOIPIDMSDPQATMGCVPNLTRGAGWPGDPVDTNPL 348  
 DB 469 NK-----EPDDE--GVMTDLMSDPDEBKGFPSK-RGIGFSPGDTITENPL 512  
 QY 349 QRRLSLVYIRSHCEKPGHGFEMHDNKKIITTFASNNYAIASNKCAVYRL-NNQMLPFPVQ 407  
 DB 513 KINMLSLIIRSHYERVDGSLGNGQLYTFSAIPNYCDINKKGAFLFKGNSIKPECVT 572  
 QY 408 Y 408  
 DB 573 F 573  
 RESULT 29  
 Q8IDF7 PLA77  
 Q8IDF7 PLA77 PRELIMINARY; PRT; 658 AA.  
 AC Q8IDF7-  
 DT 01-MAR-2003 (TREMBLrel\_23, Created)  
 DT 01-MAR-2003 (TREMBLrel\_23, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel\_26, Last annotation update)  
 DE Setine/threonine protein phosphatase pEPps (EC 3.1.3.16).  
 GN Name=pEPps; Synonyms=MAL1P1.274;  
 OS Plasmodium falciparum (isolate 3D7).  
 CC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.  
 CX NCBI\_TaxID=36329;  
 NX [1]  
 RP NCULEOTIDE SEQUENCE.  
 RC STRAIN=3D7;  
 RA Harris B., Lemard N., Clark L., Line A., Barron A., Corton C.,  
 RA Bertram M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,  
 RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrell B.;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL044509; CAD52675.1; -; Genomic DNA.  
 DR HSSP; P36873; 11T6.  
 DR GO; GO:0005737; C:cytoplasm; IEA.





[illegible]

RESULT 32	CAUFS6_THEAN	PRT;	516 AA.
ID	Q4UFS6_THEAN PRELIMINARY;		
AC	Q4UFS6;		
DT	13-SEP-2005 (TREMBLrel. 31, Created)		
DT	13-SEP-2005 (TREMBLrel. 31, last sequence update)		
DT	13-SEP-2005 (TREMBLrel. 31, last annotation update)		
DE	Serine/threonine protein phosphatase, putative (BC 3.1.3.16).		
GN	ORRNames=TA1575;		
OS	Thellieria annulata.		
OC	Eukaryota; Alveolata; Apicomplexa; Piroplasmida; Thellieridae;		
OC	Thellieria.		
OX	NCBI_TaxID=5874;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=Amkara isolate clone C9;		
RA	Pain A., Renald H., Murphy L., Harris D.A., Quail M.A., Berriman M.,		
RA	Hall N., Barrell B.G.;		
RT	"The chromosome 2 sequence of <i>Thellieria annulata</i> ."		
RL	Submitted (Apr-2005) to the EMBL/GenBank/DBD databases.		
KM	EMBL, CR940348; CA174040.1; -; Genomic_DNA.		
SD	Hydrolase.		
SEQ	SEQUENCE 516 AA; 59193 MW; C670662BEFCAPAS5 CRC64;		
Query Match	16.7%; Score 574.5; DB 2; Length 516;		
Best Local Similarity	38.2%; Pred. No. 1.6e-26;		
Matches 120; Conservative 65; Mismatches 102; Indels 27; Gaps 8			
QY	103 NHID--LIDIV--FRKKRGNELAPKRYVALILREAKSLKOLPNISPVSTAVSQOVTGQ 157		
DB	200 NHLDKSYLFPHNTLFLVILKVPAGNTHKKVYCMILFLVILKIRREYSVVDLNIYPCDELTVGQ 259		
QY	158 DLHGGLDILLVYLHNGGCPSSNPVPVQCDPFDPRKRGLEVLALLSLTLAPPAVFLNR 217		
DB	260 DIHGGFYDLLNLFSSINGEPTDSTNSLYLFNGDPVDRKSPFECVTLFLAKVLPSSFHLVR 319		
QY	218 GNHDSVNNARYGFTREYASKPRNKKHKLAFIDVRYMLPLGVSIVRSRLVHGGF--S 275		
DB	320 GNHREALNKKCGFGGELLNKK--DEKYNMLFCGSFRLPLGLYLNNKVLVIHGGLEGY 376		
QY	276 DSTSLDLIKSIDRGKTVSILRPEPLTDGEPDLKTEMQOIFDINWSDPQATMGCVPTNLRGA 335		
DB	377 ENVTIEDLKCKIDRFK-----EFSDS--GLMTDMMSDPKSNIGSLSPK-RGV 420		

[illegible]

RESULT 33	0962N7_PLAFA	PRT;	594 AA.
ID	Q962N7_PLAFA PRELIMINARY;		
AC	Q962N7;		
DT	01-DEC-2001 (TREMBLrel. 19, Created)		
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)		
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)		
DE	Serine/threonine protein phosphatase PfPp5.		
GN	Name=pp5;		
OS	Plasmodium falciparum.		
OC	Eukaryota, Alveolata, Apicomplexa, Haemosporidia, Plasmodium.		
OX	NCBI_Txid=5833;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RX	MEDLINE=21896223; PubMed=11897131; DOI=10.1016/S0166-6851(02)00007-5;		
RA	Lindenhal C., Klinkert M.O.;		
RT	"Identification and biochemical characterization of a protein		
RT	phosphatase 5 homologue from Plasmodium falciparum.";		
RL	Mol. Biochem. Parasitol. 120:257-268(2002).		
DR	EMBL, AF0404815, AAK95648.1; -; Genomic_DNA.		
DR	HSSP, P36873; 1176.		
DR	GO; GO:0005737; C:cytoplasm; IEA.		
DR	GO; GO:0005634; C:nucleus; IEA.		
DR	GO; GO:0016787; F:hydrolase activity; IEA.		
DR	GO; GO:0004721; F:phosphoprotein phosphatase activity; IEA.		
DR	GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.		
DR	InterPro; IPR004843; M-peptidase.		
DR	InterPro; IPR011236; PP2ase_5.		
DR	InterPro; IPR001440; TPR.		
DR	InterPro; IPR006186; T_phatase_apah.		
DR	Pfam; PF00149; Metallophos; 1.		
DR	Pfam; PF00515; TPR_1; 3.		
DR	PIRSP; PIRSF033096; PP2ase_5; 1.		
DR	PRINTS; PRO0114; STPHPTASE.		
DR	SMART; SMO0156; PP2ac; 1.		
DR	SMART; SMO0028; TPR; 3.		
DR	PROSITE; PSS0005; TPR; 3.		
DR	PROSITE; PSS0293; TPR_REGION; 2.		
SO	SEQUENCE 594 AA; 69248 MW; 64CA9D0004DBEEF CRC64;		

Query Match	16.7%	Score 573.5	DB 2	Length 594
Best Local Similarity	35.9%	Pred. No. 2.3e-261		
Matches 151	Conservative 77	Mismatches 144	Indels 49	Gaps 17
QY	6	IRAAIFLOKWRHRRQARREMRQRCMQLFONILEVASE--QDAELYKFENDLIKIMPOAA	63	
DB	184	IDBAIKIMPY-----AKAYRRKGSYLLISDLKRASBECFOVKLTUKONSELK-LKQCK	238	
QY	64	GR--KNYQGSANHSVLDK----DDLVBBSGDIYNALIELPI-RKHNIIDLIVFRKKR	116	
DB	239	KLIFBQCPQKAI--LEQKMPYETTLVLDLSTKEN--MEATYRRNMLNT--DPLKQVA	291	
QY	117	-----GNRLHPKYVALILREAAKSLKOLPNISPVSTAVSQOVTYCGDLHGKLDLLVL	170	
DB	292	DYISIPNNKLKKCCAYLVLDVYIKLTKELPTLVYLVNLTREDFLTICGVHQGYDILLNM	351	
QY	171	HNKGIPSSNRYVFNFGDVPDRGKGLGTEYLLLSLYLAPPAVFLPDRNGHDSVNNARYG	230	
DB	352	KINGIPBSKNSVILTFNGDVPDRGSRSEVAYILFYLAKLTFPNNVYILRGNHREDNNKAYG	411	
QY	231	FIREVBSKYPNNHKRIILAFIDEVYRMPLDGSVLSRVLIIVHGF--SDSTSLDIKSIDR	288	
DB	412	FLIGLEQECY--DEKMHVLFSDS--FEFLPLAVLVNNGINILTCHGDIPSKIDTTLBEDLEKDR	468	

QY	289	GRVYSILAPPLDGLDPLKTEWQOQ.FDIMGSPDQATMGCVPTTARAGAWRGPDTNDPL	348
DB	469	IK-----PPLDB--GVMTDILMSDPNBEKGPSPK-RGIGSFSTDTITENPL	512
QY	349	QRHRLSYVIRSRCEKPNCHGFEMHDKKITITFSASNYVAIGSKGAYIRL-NNQLMPHFVQ	407
DB	513	KINNLSTLIRSHVEADGVSLQNGOQITVPSAPYCDIMNKGAFLFKGNSITDECVT	572
QY	408	Y 408	
DB	573	F 573	
RESULT 34			
ID	Q5W6J3_ORYSA	PRELIMINARY;	483 AA.
AC	Q5W6J3;		
DT	01-FEB-2005 (TREMBLrel. 29, Created)		
DT	01-FEB-2005 (TREMBLrel. 29, Last sequence update)		
DT	01-FEB-2005 (TREMBLrel. 29, Last annotation update)		
DE	Putative serine/threonine phosphatase.		
GN	Name=OSJUNB0115P21.1;		
OS	Oryza sativa (japonica cultivar-group).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;		
OC	Ehrhartoideae; Oryzeae; Oryza.		
OK	NCBI_TaxID=39947;		
RP	[1]		
RA	NUCLEOTIDE SEQUENCE.		
RA	Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,		
RA	Chen Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,		
RA	Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,		
RA	Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C.,		
RA	Li Y.-P., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,		
RA	Wu H.-P., Shaw J.-F.;		
RL	"Oryza sativa BAC OSJUNB0115P21 genomic sequence."		
RT	Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AC135918; AAV4139.1; -; Genomic_DNA.		
DR	GO; GO:0005737; C:cytoplasm; IEA.		
DR	GO; GO:0005634; C:nucleus; IEA.		
DR	GO; GO:0016787; F:hydrolase activity; IEA.		
DR	GO; GO:0004721; F:phosphoprotein phosphatase activity; IEA.		
DR	GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.		
DR	InterPro; IPR004843; M-pesterase.		
DR	InterPro; IPR011316; PP2ase_5.		
DR	InterPro; IPR001440; TPR.		
DR	InterPro; IPR006186; T_phase_apah.		
DR	Pfam; PF00149; Metallophos; 1.		
DR	Pfam; PF00515; TPR_1; 3.		
DR	PIRSEF; PIRSF033096; PP2ase_5; 1.		
DR	PRINTS; PR00114; STPPHPTASE		
DR	Prodom; PD000252; T_phase_apah; 1.		
DR	SMART; SM00156; PP2Ac; 1.		
DR	SMART; SM00028; TPR; 3.		
DR	PROSITE; PSS0005; TPR; 3.		
DR	PROSITE; PSS0293; TPR_REGION; 1.		
DR	SEQUENCE 483 AA; 54448 MW; C652BD87A1A44DA2 CRC64;		
Query Match 16.6%; Score 570.5; DB 2; Length 483;			
Best Local Similarity 34.8%; Pred. No. 2.6e-26;			
Matches 145; Conservative 69; Mismatches 154; Indels 49; Gaps 122;			
QY	9	AIFDIKWTRRRHQARREMR-RCNMWDIPNULIEVXASQDDQAELYKFPNDILIKHMPQAAGRN	67
DB	77	ARYSKGYRRGAAYLAWGKFKALXDFQVQKRIISPND-----DATTRKUKCEKAVQ	128
QY	68	QYQSAHYSVLDDKDDLVEEFGDINAKIEL-----PIRKNHIDLIDVPRKRG	117
DB	129	KIRFEBALISVGBEGRKSVASDIDVYIIIEVEPYGPRVDGDTITIDPYKAMLDREKQK-	187
QY	118	NRLAPRYVALIIRBAKSLIKQLPNISPVSTAVSQVYTCGDIAGKLDLLVVAHKNGPS	177
DB	188	-CIRHRYAVQVYLQTVQLRSVPSSLVDVAVPDSGHFTVCAGDVHGQYFDLILNFKINGLPS	246

QY	178	SNPEVFNQDPRDREKRGLEVLLLSLTLAPNPAVFLNRGNHDSVNNARQPIFEVES	237
Db	247	ENPFLFNQDPRDGRSFSVEVLTLLPAFKCLPFTGVLTLARGNHESKSNKLTGFEGEVR	306
QY	238	KYPRNRHKLIFID---EYRWLPLGSVLNSHVLIVHG-PS-DSTSLDLKSIDRGKYV	2922
Db	307	KLGE-----AFELFAFAFCCLPFLAHVYNNVFPVHGLFSDVDGKLDIAIDR----	356
QY	293	SILRPPLTDBEPDLXTEBMOQIFDLIMWSDPQATMGCVPTLKGAGWFGPDVTNPLQRR	3522
Db	357	--FREPEPEG-----LMCEVLMSDPQPLQGRGSPK-RGVGLSFGADVTKCFIQENN	4040
QY	353	LSYVIRSHCKCNHGFPHNDKNTIFPSANYSYALGSNNGAYIRLN-NOLMPFVQY	408
Db	405	LDLVVRSHVDEGGEYIEHDGRLTVFSAPNYCDQMGNKRAYIRFTAPELKPNISSF	461
RESULT 35			
OBHIH4_LYCES			
ID	OBHIH4_LYCES	PRELIMINARY;	PRT; 485 AA.
AC	Q8HIH4;		
DT	01-MAR-2003	(TREMBLrel. 23, Created)	
DT	01-MAR-2003	(TREMBLrel. 23, Last sequence update)	
DT	01-FEB-2005	(TREMBLrel. 29, Last annotation update)	
DE	Type 5 serine/threonine phosphatase 55 kDa isoform (Type 5 protein		
DE	serine/threonine phosphatase 55 kDa isoform).		
GN	Name=PP5;		
OS	Lycopersicon esculentum (Tomato).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;		
OC	Lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.		
OX	NCBI_TaxID=4081;		
NP	(1)		
NP	NUCLEOTIDE SEQUENCE.		
RA	Pubmed:12972652; DOI=10.1104/PP.103.02617;		
RA	De la Fuente Van Bentem S., Vossen J.H., Vosseer J.B.,		
RL	Submitted (NOV-2002) to the EMBL/Genbank/DBD databases.		
DR	EMBL; AY150041; AAN64317.1; -; mRNA.		
DR	EMBL; AY182778; AAO26214.1; -; Genomic_DNA.		
DR	HSSP; P53041; 1A17.		
DR	GO; GO:0005737; C:cytoplasm; IEA.		
DR	GO; GO:0005634; C:nucleus; IEA.		
DR	GO; GO:0005486; F:binding; IEA.		
DR	GO; GO:0016787; F:hydrolase activity; IEA.		
DR	GO; GO:0004629; F:phospholipase C activity; IEA.		
DR	GO; GO:0004721; F:phosphoprotein phosphatase activity; IEA.		
DR	GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.		
DR	InterPro; IPR004843; M-pesterase.		
DR	InterPro; IPR000909; PI_PLC_X.		
DR	InterPro; IPR011236; PP2ase_5.		
DR	InterPro; IPR011440; TPR.		
DR	InterPro; IPR011990; TPR-like_helical.		
DR	InterPro; IPR006186; T_phase_apah.		
DR	Pfam; PF00149; Metallophos; 1.		
DR	Pfam; PF00515; TPR_1; 3.		
DR	PIRSF; PIRSF033096; PP2ase_5; 1.		
DR	PRINTS; PR00114; STPHPPASR.		
DR	Prodom; PD0000252; T_phase_apah; 1.		
DR	SMART; SM00156; PP2Ac; 1.		
DR	SMART; SM00028; TPR; 3.		
DR	PROSITE; PS50007; PIPLC_X_DOMAIN; 1.		
DR	PROSITE; PS50005; TPR; 3.		
DR	PROSITE; PS50293; TPR_REGION; 1.		

Seq	SEQUENCE	485 AA;	54720 MM;	FC686780B5520B68	CR664;
Query Match		16.6%;	Score 570.5;	DB 2;	length 485;
Best Local Similarity		36.1%;	Pred. No. 2.6e-26;		
Matches 140;		Conservative 55;	Mismatches 124;	Indels 69;	Gaps 12
QY	EMQRRCNNQIFQNLRYASQDQALRYKPFNDLIKMPQAAKGRKQYQSAHVSLLDDKD	83			
DB	ESQR---SVADSIDYRS-----VEVEPOYAGARIE-----	169			
QY	LYEFGDIYNAKIELPIRKNIHDLIDVERKKRGRLHPKYVALILREAKSLKOLPNS	143			
DB	170 -----GDVYTL-----DFYKMLDDPKMK--NLRKYAYQIVLQDREMLRPLSV	214			
QY	144 PVSTAVSQQVTCVGDHLGKLDLVLVLRKNGLPSSNPVFGDIPVDKRGKLEVLL	203			
DB	215 DIVPEGRKHFVTCGVHGGQFYDLLNIFELNGLPSEDNPLFLFGDIPVDGSPFLVITL	274			
QY	204 SLIYLPFAVPLNRGNHDSVNNAYGFRFVRSKYPRNHRKILFLFIDVYMLPLGSYL	263			
DB	275 AFKCMCFSAIHLARNHESKSNMKYGGFGRSKLSIFVLFPA---EVFCCLPLAVYI	331			
QY	264 NSRYLVIVGSG-PS-DSTSLDLTKSIDRGKYSILRPPLIDGSPDKTEMQGIFDIMS	321			
DB	332 NEKVYVHVGGLPSVDGKVLSDIRADR-----FCSP-----EGLMCELLMSDP	376			
QY	332 QATWGCVPNTLRGAGVWFQPDVTDNPLQHRRLSYVIRSHCKPGNHGFMDKNTITFSA	381			
DB	377 QPQPRGRPSK-RGVGLSFGQDVTKQFLQNNLDLVVRSHGVVDQVYELIENDKLITFSA	435			
QY	382 SNRYAIGSNKAYIRLN-NQLMPHFVQY 408				
DB	436 PNYCDQMGNKGAFFIREADPMKPNITVF 463				
RESULT 36					
ID	Q9BPMO_HUMAN				
AC	Q9BPMO;				
DT	01-JUN-2001 (TrEMBLrel. 17, Created)				
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)				
DT	10-MAY-2005 (TrEMBLrel. 30, Last annotation update)				
DE	PP5C protein (Fragment).				
GN	Name=PP5C;				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homi				
OC	Homio.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	NUCLEOTIDE SEQUENCE.				
RC	TISUS-Muscle;				
RX	MEDLINE=22388957; PubMed=12477932; DOI=10.1073/pnas.242603899;				
RA	Strausberg R.L., Feingold E.A., Grove L.H., Derge J.G.,				
RA	Klausner R.D., Collins F.S., Wagner L., Schenck C.M., Schuler G.D.,				
RA	Altchul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bat N.K.,				
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,				
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,				
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.R.,				
RA	Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,				
RA	Rata S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,				
RA	Boeak S.A., McGwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,				
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hults S.W.,				
RA	Villalon D.K., Mizny D.M., Sodergren E.J., Lu X., Gibbs R.A.,				
RA	Fahay J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,				
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,				
RA	Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,				
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,				
RA	Butcherfield Y.S.N., Krzyzanski M.I., Skalska U., Smallus D.E.,				
RA	Schmeh A., Schein J.E., Jones S.J.M., Marra M.A.,				
RT	"Generation and initial analysis of more than 15,000 full-length human				
RL	and mouse cDNA sequences."				
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).				

RN	[2]
RP	Nucleotide sequence.
RC	Tissue-muscle;
RA	Straubeberg R.;
RL	Submitted (Nov-2000) to the EMBL/GenBank/DBJ databases.
RN	[3]
RP	Nucleotide sequence.
RC	Tissue-muscle;
RA	Straubeberg R.;
RL	Submitted (Jan-2001) to the EMBL/GenBank/DBJ databases.
CC	-I- CATALYTIC ACTIVITY: A phosphoprotein + H(2O) = a protein + phosphate.
CC	-I- Similarity: Belongs to the PPP phosphatase family.
DR	EMBL; BC000750; AAH00750.4; -; mRNA.
DR	EMBL; BC001831; AAH01831.4; -; mRNA.
DR	HSSP; P53041; 1A17.
DR	SMR; O9BPW0; 21-481.
DR	GO; GO:0005737; Cytoplasm; IEA.
DR	GO; GO:0005634; C:nucleus; IEA.
DR	GO; GO:0005488; P:binding; IEA.
DR	GO; GO:0016787; F:hydrolyase activity; IEA.
DR	GO; GO:0004721; F:phosphoprotein phosphatase activity; IEA.
DR	GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR	InterPro; IPR004643; M-pesterase.
DR	InterPro; IPR011261; PPtase_5.
DR	InterPro; IPR006186; T_phctase_aprh.
DR	InterPro; IPR011990; TPR-like_helical.
DR	Pfam; PF001149; Metallophos; 1.
DR	Pfam; PF00515; TPR_1; 3.
DR	PIRSF; PIRSF033096; ppptase_5; 1.
DR	PRINTS; PR00114; STEPHPTASE.
DR	Prodrom; PD000252; T_phctase_aprh; 1.
DR	SMART; SMO0156; PP2Ac; 1.
DR	SMART; SMO0028; TPR; 3.
DR	PROSITE; PS00125; SER_THR_PHOSPHATASE; UNKNOWN_1.
DR	PROSITE; PS50005; TPR; 3.
DR	PROSITE; PS50293; TPR_REGION; 1.
KW	Hydrolase; Iron; TPR repeat.
FT	NON TER
SO	SEQUENCE 484 AA; 55096 MW; 152A9C37AE3D4F69 CRC64;
QY	Query Match 16.6%; score 570; DB 2; Length 484;
BEST	Local Similarity 33.7%; Pred. No. 2.8e-26;
MATCHES	143; Conservative 71; Mismatches 142; Indels 68; Gaps 16
OY	11 FIQKWTTRHQARRMORRRCMQITONTLEYASBOQAELYEKFNDLIKMPPOAGAKNQYO 70 :: Db 93 YIKGYERRAAASNMAIGK-----FR---AALRD-----YETVKKPKPHDKAKMKYO 135
OY	71 -----GSAHY-SVLDDKD-----DLVERFG--DIYNATEILPIRKNHIDL 107
Dy	136 ECKKIYOKAFERAILADGEHKRSVAVDLSIDISMTTEDYSGRPKLDGVTTTIFPMVE--- 191
OY	108 LIDVYRKKGKRNLAPKYVALILIREAAKLKOLPNISPSTAVSQCVTVCGDLHGKLDLIL 167 :: Db 192 LMGWYKOQG-KLRKKAQYQLLVQVKYSLSTLTVELTKTEKITVTCGDTHNGFYDLL 249
OY	168 VTLHKNGLPSSSNRYVFNQDPVDNRGKGLLEVLILLISLYLAFPNAVFPLRNQHEDSVNNA 227 :: Db 250 NIPFLNGLPSRETNPFIYIFWGDFVDDRGSFSVREYILTLLFGKLLYPDHFFHLRGNHKEFDNMNO 309
OY	228 RYGRTREVESKYPPNNHRIILAFLDEYVRWLPLSGVLNSRVYLIVGGCF--SDSTSILDLS 285 :: Db 310 IYGBEGEYKAYKT--ADMVELFSEVFPEMLPLAACINKNVLMHGGLFSBDSGVTLDDLIRK 366
OY	286 IDRKYSVIARPLLTDGSPFLTDTQMOLFIDIMMSDPOATMCVPRTLIRGAQVWFQPDVTD 345 :: Db 367 IERRK-----QPF--DGSPM-----CDLMDSDPOANG-RATSISKVGSCGFQPDVTX 410
OY	346 NFLQRHRLSYVIRSHCEKPNGHFMHDNKIIITTFISASYIAIGSKNCAVYIRL-NQOLMPH 404
DB	411 AFLERNMLDYTIIRSHFEVVAAGVGRCTVTSAPYPCOMGNKASYIHLOGSDLRPO 470



QY 405 FVOY 408  
Db 471 FHOF 474

RESULT 37  
PP5\_HUMAN STANDARD; PRT; 499 AA.  
ID PP5\_HUMAN  
AC P53041; Q16722;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Serine/threonine protein phosphatase 5 (EC 3.1.3.16) (PP5) (Protein  
phosphatase T) (PP-T) (PPT).  
GN Name=PP5C; Synonym=PP5;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Barchonoglires; Primates; Catarrhini; Homidae;  
OC Homo.  
NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE OF 7-499.  
RX MEDLINE=9500929; PubMed=7925273;  
RA Chen M.X., McPartlin A.E., Brown L., Chen Y.H., Barker H.M.,  
Cohen P.T.W.;  
RT "A novel human protein serine/threonine phosphatase, which possesses  
four tetratricopeptide repeat motifs and localizes to the nucleus.";  
RL EMBO J. 13:4278-4290(1994).  
RN [2]  
RP NUCLEOTIDE SEQUENCE OF 9-499.  
RX MEDLINE=96115607; PubMed=8666404;  
RA Yong W.H., Ueki K., Chou D., Reeves S.A., von Deimling A.,  
Gusella J.F., Mohrenweiser H.W., Buckler A.J., Louis D.N.;  
RT "Cloning of a highly conserved human protein serine-threonine  
phosphatase gene from the glioma candidate region on chromosome  
19q13.3.";  
RL Genomics 29:533-536(1995).  
RN [3]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RX PubMed=15057824; DOI=10.1038/nature02399;  
RA Grimwood J., Gordon L.A., Olsen A.S., Terry A., Schmutz J.,  
Lamerdin J.E., Hellsten U., Goodstein D., Conrume O., Tran-Gyamfi M.,  
Aerts A., Altherr M., Ashworth L., Bajorek E., Black S., Branncomb E.,  
Caenepeel S., Carrano A.V., Cawley C., Chan Y.M., Christensen M.,  
Cleland C.A., Copeland A., Dalin B., Dehal P., Denye M., Dettler J.C.,  
Bacopa J., Flowers D., Fotopulos D., Garcia C., Georgescu A.M.,  
Glavina T., Gomez M., Gonzales B., Groza M., Hammon N., Hawkins T.,  
Haydu L., Ho I., Huang W., Israni S., Jett J., Kadner K., Kimball H.,  
Kobayashi A., Laitonov V., Leem S.-H., Lopez F., Lou Y., Lowry S.,  
Malafait S., Martinez D., McCreedy P.M., Medina C., Morgan J.,  
Nelson K., Nolan M., Ovcharenko I., Pittluck S., Pollard M.,  
Popkie A.P., Predki P., Quan G., Ramirez L., Raab S., Retterer J.,  
Rodriguez A., Rogers S., Salamov A., Salazar A., She X., Smith D.,  
Slezak T., Solovyev V., Thayer N., Tice H., Tsai M., Ustaszewska A.,  
Vo N., Wagner M., Wheeler J., Wu K., Xie G., Xie J., Yang J., Dubchak I.,  
Pruett T.S., DeLong P., Dickson M., Gordon D., Richter E.B.,  
Pennacchio L.A., Richardson P., Stubbs L., Rokhsar D.S., Myers R.M.,  
Rubin E.M., Lucas S.M.,  
RT "The DNA sequence and biology of human chromosome 19.";  
RL Nature 428:529-535(2004).  
RN [4]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Peingold B.A., Grouse L.H., Derge J.G.,  
Klauser R.D., Collins F.S., Wagner C.M., Schuler G.D.,  
Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,  
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
Datchenko L., Marusik K., Farmer A., Rubin G.M., Hong L.,  
Stapleton M., Soares M.B., Bonaldo M.F., Cassavant T.L., Scheetz T.E.,  
Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
Bosak S.A., McSwan P.J., McKernan K.U., Malek J.A., Gunaratne P.H.,  
Richard S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
Villalon D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,  
Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,  
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Smailus D.B.,  
Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smailus D.B.,  
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [5]  
RP NUCLEOTIDE SEQUENCE OF 1-37.  
RC TISSUE=Fetal brain;  
RX MEDLINE=96144708; PubMed=8561788; DOI=10.1006/bbrc.1996.0092;  
RA Xu X., Lagercrantz J., Zickert P., Bajalica-Lagercrantz S.,  
Zetterberg A.;  
RT "Chromosomal localization and 5' sequence of the human protein  
serine/threonine phosphatase 5' gene.";  
RL Biochem. Biophys. Res. Commun. 218:514-517(1996).  
RN [6]  
RP INTERACTIONS WITH CDC16 AND CDC27.  
RX MEDLINE=98070360; PubMed=9405394; DOI=10.1074/jbc.272.51.32011;  
RA Ollendorff V., Donoghue D.J.;  
RT "The serine/threonine phosphatase PP5 interacts with CDC16 and CDC27,  
two tetratricopeptide repeat-containing subunits of the anaphase-  
promoting complex.";  
RL J. Biol. Chem. 272:32011-32018(1997).  
RN [7]  
RP X-RAY CRYSTALLOGRAPHY (2.45 ANGSTROMS) OF 19-177.  
RX MEDLINE=98151343; PubMed=9482716; DOI=10.1093/emboj/17.5.1192;  
RA Das A.K., Cohen P.W., Barford D.;  
RT "The structure of the tetratricopeptide repeats of protein phosphatase  
5: implications for TPR-mediated protein-protein interactions.";  
RL EMBO J. 17:1192-1199(1998).  
CC -1- FUNCTION: May play a role in the regulation of RNA biogenesis  
and/or mitosis. In vitro, dephosphorylates serine residues of  
skeletal muscle phosphorylase and histone H1.  
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +  
phosphate.  
CC -1- COFACTOR: Binds 1 iron ion per subunit (By similarity).  
CC -1- COFACTOR: Binds 1 manganese ion per subunit (By similarity).  
CC -1- SUBUNIT: Interacts with CDC16 and CDC27.  
CC -1- SUBCELLULAR LOCATION: Nuclear; predominantly, but also present in  
the cytoplasm.  
CC -1- TISSUE SPECIFICITY: Ubiquitous.  
CC -1- SIMILARITY: Belongs to the PPP phosphatase family. PP-T subfamily.  
CC -1- SIMILARITY: Contains 3 TPR repeats.  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use as long as its content is in no way modified and this statement is not  
removed.  
DR EMBL; X89416; CAA61595.1; -; mRNA.  
DR EMBL; U25174; AAB60384.1; -; mRNA.  
DR EMBL; AC007193; AAD22669.1; -; Genomic DNA.  
DR EMBL; BC001970; AA01970.1; -; mRNA.  
DR EMBL; X91211; CAA63089.1; -; mRNA.  
DR PIR; S52570; S52570.  
DR PDB; 1A17; X-ray; @=16-181.  
DR PDB; 1S95; X-ray; A/B=169-499.  
DR SMR; P53041; 23-499.  
DR Ensembl; ENSG0000011485; Homo sapiens.  
DR HGN; HGN:9322; PP5C.  
DR H-InvDB; HIX015250; -.  
DR MIM; 600658; -.  
DR GO; GO:0005634; C:nucleus; TAS.  
DR GO; GO:004722; F:protein serine/threonine phosphatase activity; TAS.

```

DR GO; GO:0004871; P:signal transducer activity; IMP.
DR GO; GO:0007067; P:mitosis; TAS.
DR GO; GO:0043123; P:protein regulation of I-kappaB kinase/NF-kappaB; IMP.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.
DR GO; GO:0006350; P:transcription; TAS.
DR InterPro; IPR004843; M-pesterase.
DR InterPro; IPR011236; P:peptase_5.
DR InterPro; IPR006186; T:phatase_apah.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR011990; TPR-like_helical.
DR Pfam; PF00149; Metallophos; 1.
DR Pfam; PF00515; TPR_1; 3.
DR PIRSF; PIRSF03096; peptase_5; 1.
DR PRINTS; PR00114; STEPHPTAS.
DR ProDom; PD000252; T_phatase_apah; 1.
DR SMART; SM00156; PP2AC; 1.
DR SMART; SM00028; TPR; 3.
DR PROSITE; PS00125; SER_THR_PHOSPHATASE; 1.
DR PROSITE; PS50005; TPR; 3.
DR PROSITE; PS50293; TPR_REGION; 1.
DR 3D-structure; Hydrolyase; Iron; Manganese; Metal-binding;
KM Nuclear protein; Protein phosphatase; Repeat; TPR repeat.
FT REPEAT 28 61
FT REPEAT 62 95
FT REPEAT 96 129
FT REGION 184 499
FT ACT SITE 304 304
FT METAL 242 242
FT METAL 244 244
FT METAL 271 271
FT METAL 303 303
FT METAL 352 352
FT METAL 427 427
FT CONFLICT 403 403
FT HELIX 22 40
FT TURN 41 42
FT HELIX 44 57
FT TURN 59 60
FT HELIX 62 74
FT TURN 75 76
FT HELIX 78 91
FT TURN 93 94
FT HELIX 96 108
FT TURN 109 110
FT HELIX 112 125
FT TURN 127 128
FT HELIX 130 164
FT TURN 168 169
SQ SEQUENCE 499 AA; 56879 MW; DB3B2090D8658BB3 CRC64;

Query Match 16.6%; Score 570; DB 1; Length 499;
Best Local Similarity 33.7%; Pred. No. 2, 9e-26;
Matches 143; Conservative 71; Mismatches 142; Indels 68; Gaps 16;

11 FIDKWRRRHQAREREMORRCNMWIFONILEVASBODQAEVYKFNDLIKMPQAAGRNOGO 70
Db 95 YIKGYRRAASNMALGK-----FR-----AALRD-----YETVAVKAPHDKDAKMKQ 137
71 -----GSAHV-SVLDDKD-----DLVEERG-DIVAKIELPIRKNHIDL 107
138 ECKNIYKQAFERAIAGDEHKRSVVDLSIDSWTIDERSGRLDGKVTISPMKE----- 193
108 LDVDFKKKGKGRNLRHPYVALILREAAKSLKQLPNISPVSTAVSQVTVCGDLHGKLDLIL 167
194 LMQWYDQK-KLHRCAYQILVQVKEVLSKSLTVEITLKEKELTVCGDTHGQFYDL 251
168 VVLHKGGLPSSSNPYVNGDFVDRGKRGLEVLILLSLYAFNNAVFLNRGNHDSVMA 227
252 NIEELNGLSSETNPYIFNEDFVDRGSPSEVILLTGFKLLYDHFHLRGNHETDMNQ 311
228 RYGFIREVSKYPRNKRILAFIDVYRMLPLGSLVNSRYLVHGGF--SDSTSLDLIS 285
312 IYGFBSGVAKYK---AQWYELFSVFEMLPFQAQINGKVLIMHGGLFSEDDVTLDDIRK 368

```

```

QY 286 IDRGKTVSTLRPEPLDGEPLDKTEWQIFDLMNSDPQATMGCVPTNLGAGVIRGPDVTD 345
Db 369 IERNK-----QPP--DSGPM-----CDLLMSDPQPNQ-RSISKGVSCQFEPDVTX 412
QY 346 NFIQRRLSYVIRSHCKPKNHGFPMHDKIIRIFASNYVYAIGSKGAYIRL-NNQLMPH 404
Db 413 ALLENNLDYIIRSHVYAEAGVVAHGRCVTFSPAPNYCDQMGKASTYIHLQSGDLRPQ 472
QY 405 FVOY 408
Db 473 FHOY 476

RESULT 38
Q53FR0_HUMAN
ID Q53FR0_HUMAN PRELIMINARY; PRT; 499 AA.
AC Q53FR0;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Protein phosphatase 5, catalytic subunit variant (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea;
OC Homo.
OC NCBI_Taxid=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=stomach mucosa;
RA Maruyama K., Sugano S.;
RT "oligo-capping : a simple method to replace the cap structure of
RT eucaryotic mRNAs with oligoribonucleotides.";
RL Gene 138:171-174(1994).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=stomach mucosa;
RA Suzuki Y., Yoshitomo K., Maruyama K., Sugama A., Sugano S.;
RT "Construction and characterization of a full length-enriched and a 5'-
RT end-enriched cDNA library.";
RL Gene 200:149-156(1997).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=stomach mucosa;
RA Suzuki Y., Sugano S., Totoki Y., Toyoda A., Takeda T., Sakaki Y.,
RA Tanaka A., Yokoyama S.;
RL Submitted (Apr-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -!- SIMILARITY: Belongs to the PPP phosphatase family.
DR EMBL; AK232222; BAD96942.1; -, mRNA.
DR SMR; Q53FR0; 23-499.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0004720; F:phosphoprotein phosphatase activity; IEA.
DR GO; GO:0006470; F:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR004843; M-pesterase.
DR InterPro; IPR006186; T_phatase_apah.
DR InterPro; IPR001440; TPR.
DR Pfam; PF00149; Metallophos; 1.
DR Pfam; PF00515; TPR_1; 3.
DR PRINTS; PR00114; STEPHPTAS.
DR ProDom; PD000252; T_phatase_apah; 1.
DR SMART; SM00156; PP2AC; 1.
DR SMART; SM00028; TPR; 3.
DR PROSITE; PS00125; SER_THR_PHOSPHATASE; UNKNOWN_1.
DR PROSITE; PS50005; TPR_3.
DR PROSITE; PS50293; TPR_REGION; 1.
DR 3D-structure; Iron; TPR repeat.
FT NON TER 1 1
SQ SEQUENCE 499 AA; 56907 MW; DB5E20D85465C3F CRC64;

```



